

# An introduction to Lie group integrators – basics, new developments and applications

Elena Celledoni

Håkon Marthinsen

Brynjulf Owren

July 3, 2012

## Abstract

We give a short and elementary introduction to Lie group methods. A family of symplectic integrators on cotangent bundles of Lie groups is presented and the notion of discrete gradient methods is generalized to Lie groups. Finally, a selection of applications of Lie group integrators are discussed.

## 1 Introduction

The significance of the geometry of differential equations was well understood already in the nineteenth century, and in the last few decades such aspects have played an increasing role in numerical methods for differential equations. Nowadays, there is a rich selection of integrators which preserve properties like symplecticity, reversibility, phase volume and first integrals, either exactly or approximately over long times [28]. Differential equations are inherently connected to Lie groups, and in fact one often sees applications in which the phase space is a Lie group or a manifold with a Lie group action. In the early nineties, two important papers appeared which used the Lie group structure directly as a building block in the numerical methods. Crouch and Grossman [20] suggested to advance the numerical solution by computing flows of vector fields in some Lie algebra. Lewis and Simo [42] wrote an influential paper on the Lie group based integrators for Hamiltonian problems, considering the preservation of symplecticity, momentum and energy. These ideas were developed in a systematic way throughout the nineties by several authors. In a series of three papers, Munthe-Kaas [50, 51, 52] presented what are now known as the Runge–Kutta–Munthe-Kaas methods. By the turn of the millennium, a survey paper [33] summarized most of what was known by then about Lie group integrators.

The purpose of the present paper is three-fold. First, in section 2 we would like to give an elementary, geometric introduction to the ideas behind Lie group integrators. Secondly, we present some new material in sections 3 and 4. Symplectic Lie group integrators have been known for some time, derived by Marsden and coauthors [45] by means of variational principles. Here we consider a group structure on the cotangent bundle of a Lie group

and derive symplectic Lie group integrators using the model for vector fields on manifolds defined by Munthe-Kaas in [52]. In section 4 we extend the notion of discrete gradient methods as proposed by Gonzalez [27] to Lie groups, and thereby we obtain a general method for preserving first integrals in differential equations on Lie groups. Our third objective is to present a selection of applications of Lie group integrators. There are many such examples to choose from, and we give here only a few teasers.

We would also like to briefly mention some of the issues we are *not* pursuing in this article. One is the important family of Lie group integrators for problems of linear type, including methods based on the Magnus and Fer expansions. An excellent review of the history, theory and applications of such integrators can be found in [2]. We will also skip all discussions of order analysis of Lie group integrators. This is a large area by itself which involves technical tools and mathematical theory which we do not wish to include in this relatively elementary exposition. There have been several new developments in this area recently, in particular by Lundervold and Munthe-Kaas, see e.g. [44].

## 2 Lie group integrators

The simplest consistent method for solving ordinary differential equations is the Euler method. For an initial value problem of the form

$$\dot{y} = F(y), \quad y(0) = y_0,$$

one takes a small time increment  $h$ , and approximates  $y(h)$  by the simple formula

$$y_1 = y_0 + hF(y_0),$$

advancing along the straight line coinciding with the tangent at  $y_0$ . Another way of thinking about the Euler method is to consider the constant vector field  $F_{y_0}(y) := F(y_0)$  obtained by parallel translating the vector  $F(y_0)$  to all points of phase space. A step of the Euler method is nothing else than computing the exact  $h$ -flow of this simple vector field starting at  $y_0$ . In Lie group integrators, the same principle is used, but allowing for more advanced vector fields than the constant ones. A Lie group generalization of the Euler method is called the Lie–Euler method, and we shall illustrate its use through an example [20].

**Example 1, the Duffing equation.** Consider the system in  $\mathbf{R}^2$

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -ax - bx^3, \end{aligned} \quad a \geq 0, b \geq 0, \tag{1}$$

a model used to describe the buckling of an elastic beam. Locally, near a point  $(x_0, y_0)$  we could use the approximate system

$$\begin{aligned} \dot{x} &= y, & x(0) &= x_0, \\ \dot{y} &= -(a + bx_0^2)x, & y(0) &= y_0, \end{aligned} \tag{2}$$

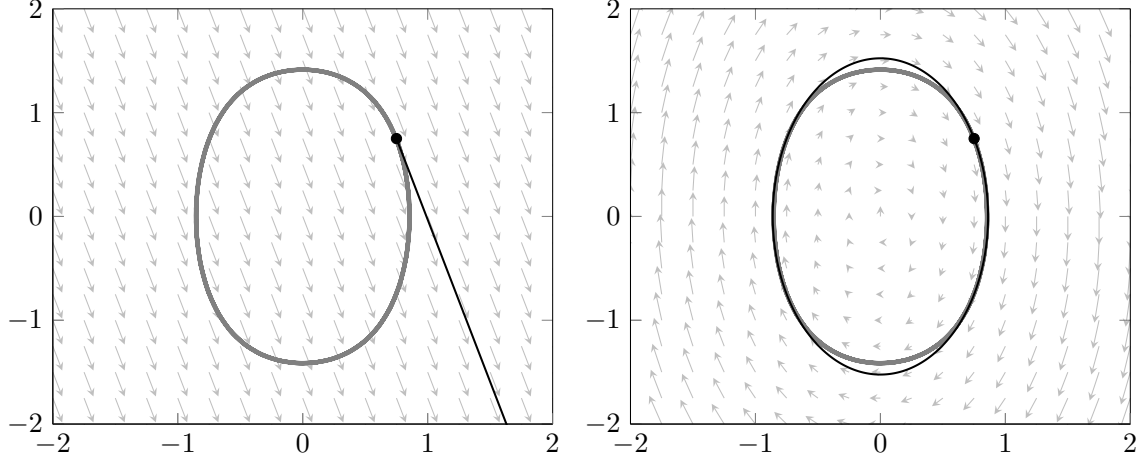


Figure 1:  $(\mathbf{R}^d, +)$ -frozen vector field (left) and  $SL(2)$ -frozen vector field (right) for the Duffing equation

which has the exact solution

$$\bar{x}(t) = x_0 \cos \omega t + \frac{y_0}{\omega} \sin \omega t, \quad \bar{y}(t) = y_0 \cos \omega t - \omega x_0 \sin \omega t, \quad \omega = \sqrt{a + bx_0^2}.$$

Alternatively, we may consider the local problem

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= -ax - bx_0^3, \end{aligned} \tag{3}$$

having exact solution

$$\begin{aligned} \bar{x}(t) &= x_0 \cos \alpha t + \frac{y_0}{\alpha} \sin \alpha t + bx_0^3 \frac{\cos \alpha t - 1}{\alpha^2}, \\ \bar{y}(t) &= y_0 \cos \alpha t - \alpha x_0 \sin \alpha t - bx_0^3 \frac{\sin \alpha t}{\alpha}, \end{aligned} \quad \alpha = \sqrt{a}. \tag{4}$$

In each of the two cases, one may take  $x_1 = \bar{x}(h)$ ,  $y_1 = \bar{y}(h)$  as the numerical approximation at time  $t = h$ . The same procedure is repeated in subsequent steps. A common framework for discussing these two cases is provided by the use of frames, i.e. a set of vector fields which at each point is spanning the tangent space. In the first case, the numerical method applies the frame

$$X = \begin{bmatrix} y \\ 0 \end{bmatrix} =: y \partial x, \quad Y = \begin{bmatrix} 0 \\ x \end{bmatrix} =: x \partial y. \tag{5}$$

Taking the negative Jacobi–Lie bracket between  $X$  and  $Y$  yields the third element of the standard basis for the Lie algebra  $\mathfrak{sl}(2)$ , i.e.

$$H = -[X, Y] = x \partial x - y \partial y, \tag{6}$$

so that the frame may be augmented to consist of  $\{X, Y, H\}$ . In the second case, the vector fields  $E_1 = y \partial x - ax \partial y$  and  $E_2 = \partial y$  can be used as a frame, but again we choose to augment these two fields with the commutator  $E_3 = -[E_1, E_2] = \partial x$  to obtain the Lie algebra of the special Euclidean group  $SE(2)$  consisting of translations and rotations in the plane.

In general, a way to think about Lie group integrators is that we have a manifold  $M$  where there is such a frame available;  $\{E_1, \dots, E_d\}$  such that at any point  $p \in M$  one has

$$\text{span}\{E_1(p), \dots, E_d(p)\} = T_p M$$

and frames with this property are said to be locally transitive. This may be a linear space or in many cases even a Lie algebra of vector fields. In the example with Duffing's equation, the set  $\{X, Y, H\}$  is locally transitive on  $\mathbf{R}^2 \setminus \{0\}$  and  $\{E_1, E_2, E_3\}$  is locally transitive on  $\mathbf{R}^2$ .

Given an arbitrary vector field  $F$  on  $M$ , then at any point  $p \in M$  there exists a vector field  $F_p$  in the span of the frame vector fields such that  $F_p(p) = F(p)$ . An explicit way of writing this is by using a set of basis vector fields  $E_1, \dots, E_d$  for  $\mathfrak{g}$ , such that any smooth vector field  $F$  has a representation

$$F(y) = \sum_{k=1}^d f_k(y) E_k(y), \quad (7)$$

for some functions  $f_k: M \rightarrow \mathbf{R}$ . The vector fields  $F_p \in \mathfrak{g}$ , called *vector fields with frozen coefficients* by Crouch and Grossman [20], are then obtained as

$$F_p(y) = \sum_{k=1}^d f_k(p) E_k(y).$$

In the example with the Duffing equation we took  $E_1 = X, E_2 = Y, f_1(x, y) = 1$  and  $f_2(x, y) = -(a + bx^2)$ . The Lie–Euler method reads in general

$$y_{n+1} = \exp(hF_{y_n})y_n,$$

where  $\exp$  denotes the flow of a vector field.

A more interesting example, also found in [20] is obtained by choosing  $M = S^2$ , the 2-sphere. A suitable way to induce movements of the sphere is that of rotations, that is, by introducing the Lie group  $SO(3)$  consisting of orthogonal matrices with unit determinant. The corresponding Lie algebra  $\mathfrak{so}(3)$  of vector fields are spanned by

$$E_1(x, y, z) = -z \partial y + y \partial z, \quad E_2(x, y, z) = z \partial x - x \partial z, \quad E_3(x, y, z) = -y \partial x + x \partial y.$$

We note that  $x E_1(x, y, z) + y E_2(x, y, z) + z E_3(x, y, z) = 0$ , showing that the functions  $f_k$  in (7) are not unique. A famous example of a system whose solution evolves on  $S^2$  is the

free rigid body Euler equations

$$\dot{x} = \left(\frac{1}{I_3} - \frac{1}{I_2}\right)yz, \quad \dot{y} = \left(\frac{1}{I_1} - \frac{1}{I_3}\right)xz, \quad \dot{z} = \left(\frac{1}{I_2} - \frac{1}{I_1}\right)xy, \quad (8)$$

where  $x, y, z$  are the coordinates of the angular momentum relative to the body, and  $I_1, I_2, I_3$  are the principal moments of inertia. A choice of representation (7) is obtained with

$$f_1(x, y, z) = -\frac{x}{I_1}, \quad f_2(x, y, z) = -\frac{y}{I_2}, \quad f_3(x, y, z) = -\frac{z}{I_3},$$

so that the ODE vector field can be expressed in the form

$$F(x, y, z) = -\frac{x}{I_1} \begin{bmatrix} 0 \\ -z \\ y \end{bmatrix} - \frac{y}{I_2} \begin{bmatrix} z \\ 0 \\ -x \end{bmatrix} - \frac{z}{I_3} \begin{bmatrix} -y \\ x \\ 0 \end{bmatrix}.$$

We compute the vector field with coefficients frozen at  $p_0 = (x_0, y_0, z_0)$ ,

$$F_{p_0}(x, y, z) = \begin{bmatrix} 0 & \frac{z_0}{I_3} & -\frac{y_0}{I_2} \\ -\frac{z_0}{I_3} & 0 & \frac{x_0}{I_1} \\ \frac{y_0}{I_2} & -\frac{x_0}{I_1} & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

The  $h$ -flow of this vector field is the solution of a linear system of ODEs and can be expressed in terms of the matrix exponential  $\expm(hF_{p_0})$ . The Lie–Euler method can be expressed as follows:

Given  $p_0 = (x_0, y_0, z_0)$   
for  $n = 0, 1, \dots$   
 $p_{n+1} = \expm(hF_{p_n}) \cdot p_n$   
end for

Notice that the matrix to be exponentiated belongs to the matrix group  $\mathfrak{so}(3)$  of real skew-symmetric matrices. The celebrated Rodrigues' formula

$$\expm(A) = I + \frac{\sin \alpha}{\alpha} A + \frac{1 - \cos \alpha}{\alpha^2} A^2, \quad \alpha^2 = \|A\|_2^2 = \frac{1}{2} \|A\|_F^2, \quad A \in \mathfrak{so}(3),$$

provides an inexpensive way to compute this.

Whereas the notion of frames was used by Crouch and Grossman in their pioneering work [20], a different type of notation was used in a series of papers by Munthe-Kaas [50, 51, 52], see also [44] for a more modern treatment. Let  $G$  be a finite dimensional Lie group acting transitively on a manifold  $M$ . This means that for any two points  $m_1, m_2 \in M$

there exists a group element  $g \in G$  such that  $m_2 = g \cdot m_1$ . We denote the Lie algebra of  $G$  by  $\mathfrak{g}$ . For any element  $\xi \in \mathfrak{g}$  there exists a vector field on  $M$

$$X_\xi(m) = \left. \frac{d}{dt} \right|_{t=0} \exp(t\xi) \cdot m =: \lambda_*(\xi)(m) \quad (9)$$

Munthe-Kaas introduced a generic representation of a vector field  $F \in \mathcal{X}(M)$  by a map  $f: M \rightarrow \mathfrak{g}$  such that

$$F(m) = \lambda_*(f(m))(m) \quad (10)$$

The corresponding frame is obtained as  $E_i = \lambda_*(e_i)$  where  $e_1, \dots, e_d$  is some basis for  $\mathfrak{g}$  and one chooses the functions  $f_i: M \rightarrow \mathbf{R}$  such that  $f(m) = \sum_{i=1}^d f_i(m)e_i$ . The map  $\lambda_*$  is an anti-homomorphism of the Lie algebra  $\mathfrak{g}$  into the Lie algebra of vector fields  $\mathcal{X}(M)$  under the Jacobi–Lie bracket, meaning that

$$\lambda_*([X_m, Y_m]_{\mathfrak{g}}) = -[\lambda_*(X_m), \lambda_*(Y_m)]_{\text{JL}}$$

This separation of the Lie algebra  $\mathfrak{g}$  from the manifold  $M$  allows for more flexibility in the way we represent the frame vector fields. For instance, in the example with Duffing’s equation and the use of  $\mathfrak{sl}(2)$ , we could have used the matrix Lie algebra with basis elements

$$X_m = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Y_m = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad H_m = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

rather than the basis of vector fields (5), (6). The group action by  $g \in SL(2)$  on a point  $m \in \mathbf{R}^2$  would then be simply  $g \cdot m$ , matrix-vector multiplication, and the exp in (9) would be the matrix exponential. The map  $f(x, y)$  would in our case be

$$f: (x, y) \mapsto \begin{bmatrix} 0 & y \\ -(a + bx^2) & 0 \end{bmatrix}$$

but note that since the dimension of the manifold is just two whereas the dimension of  $\mathfrak{sl}(2)$  is three, there is freedom in the choice of  $f$ . In the example we chose not to use the third basis element  $H$ .

## 2.1 Generalizing Runge–Kutta methods

In order to construct general schemes, as for instance a Lie group version of the Runge–Kutta methods, one needs to introduce intermediate stage values. This can be achieved in a number of different ways. They all have in common that when the methods are applied in the Euclidean space where the Lie group is  $(\mathbf{R}^m, +)$ , they reduce to conventional Runge–Kutta schemes. Let us begin by studying the simple second order Heun method, sometimes called the improved Euler method.

$$k_1 = F(y_n), \quad k_2 = F(y_n + hk_1), \quad y_{n+1} = y_n + \frac{1}{2}h(k_1 + k_2)$$

Geometrically, we may think of  $k_1$  and  $k_2$  as constant vector fields, coinciding with the exact ODE  $F(y)$  at the points  $y_n$  and  $y_n + hk_1$  respectively. The update  $y_{n+1}$  can be interpreted at least in three different ways,

$$\exp\left(\frac{h}{2}(k_1 + k_2)\right)y_n, \quad \exp\left(\frac{h}{2}k_1\right) \circ \exp\left(\frac{h}{2}k_2\right)y_n, \quad \exp\left(\frac{h}{2}k_2\right) \circ \exp\left(\frac{h}{2}k_1\right)y_n \quad (11)$$

The first is an example of a Runge–Kutta–Munthe-Kaas method and the second is an example of a Crouch–Grossman method. All three fit into the framework of *commutator-free* Lie group methods. All three suggestions above are generalizations that will reduce to Heun’s method in  $(\mathbf{R}^m+)$ , in principle we could extend the idea to Runge–Kutta methods with several stages

$$y_{n+1} = y_n + h \sum_{i=1}^s b_i F(Y_i), \quad Y_i = y_n + h \sum_{j=1}^s a_{ij} F(Y_j), \quad i = 1, \dots, s,$$

by for instance interpreting the summed expressions as vector fields with frozen coefficients whose flows we apply to the point  $y_n \in M$ . But it is unfortunately not true that one in this way will retain the order of the Runge–Kutta method when applied to cases where the acting group is nonabelian.

Let us first describe methods as proposed by Munthe-Kaas [52], where one may think of the method simply as a change of variable. As before, we assume that the action of  $G$  on  $M$  is locally transitive. Since the exponential mapping is a local diffeomorphism in some open set containing  $0 \in \mathfrak{g}$ , it is possible to represent *any* smooth curve  $y(t)$  on  $M$  in some neighborhood of a point  $p \in M$  by means of a curve  $\sigma(t)$  through the origin of  $\mathfrak{g}$  as follows

$$y(t) = \exp(\sigma(t)) \cdot p, \quad \sigma(0) = 0, \quad (12)$$

though  $\sigma(t)$  is not necessarily unique. We may differentiate this curve with respect to  $t$  to obtain,

$$\dot{y}(t) = \lambda_*(\text{dexp}_{\sigma(t)} \dot{\sigma}(t))(y(t)) = F(y(t)) = \lambda_*(f(\exp(\sigma(t)) \cdot p)(y(t))) \quad (13)$$

The details are given in [52] and the map  $\text{dexp}_\sigma: \mathfrak{g} \rightarrow \mathfrak{g}$  was derived by Hausdorff in [30] as an infinite series of commutators

$$\text{dexp}_\sigma(v) = v + \frac{1}{2}[\sigma, v] + \frac{1}{6}[\sigma, [\sigma, v]] + \dots = \sum_{k=0}^{\infty} \frac{1}{(k+1)!} \text{ad}_\sigma^k v = \left. \frac{\exp(z) - 1}{z} \right|_{z=\text{ad}_\sigma} v \quad (14)$$

with the usual definition of  $\text{ad}_u(v)$  as the commutator  $[u, v]$ . The map  $\lambda_*$  does not have to be injective, but a sufficient condition for (13) to hold is that

$$\dot{\sigma} = \text{dexp}_\sigma^{-1}(f(\exp(\sigma) \cdot p))$$

This is a differential equation for  $\sigma(t)$  on a linear space, and one may choose any conventional integrator for solving it. The map  $\text{dexp}_\sigma^{-1}: \mathfrak{g} \rightarrow \mathfrak{g}$  is the inverse of  $\text{dexp}_\sigma$  and can

also be interpreted as the right trivialized derivative of the logarithm. From (14) we find that one can write  $\text{dexp}_\sigma^{-1}(v)$  as

$$\text{dexp}_\sigma^{-1}(v) = \frac{z}{\exp(z) - 1} \Big|_{z=\text{ad}_\sigma} v = v - \frac{1}{2}[\sigma, v] + \frac{1}{12}[\sigma, [\sigma, v]] + \cdots \quad (15)$$

The coefficients appearing in this expansion are scaled Bernoulli numbers  $\frac{B_k}{k!}$ , and  $B_{2k+1} = 0$  for all  $k \geq 1$ . One step of the resulting Runge–Kutta–Munthe-Kaas method is then expressed in terms of evaluations of the map  $f$  as follows

$$y_1 = \exp\left(h \sum_{i=1}^s b_i k_i\right) \cdot y_0$$

$$k_i = \text{dexp}_{h \sum_j a_{ij} k_j}^{-1} f \left( \exp\left(h \sum_j a_{ij} k_j\right) \cdot y_0 \right), \quad i = 1, \dots, s.$$

This is not so surprising seen from the perspective of the first alternative in (11), the main difference is that the stages  $k_i$  corresponding to the frozen vector fields  $\lambda_*(k_i)$  need to be “corrected” by the  $\text{dexp}^{-1}$  map. Including this map in computational algorithms may seem awkward, however, fortunately truncated versions of (15) may frequently be used. In fact, by applying some clever tricks involving graded free Lie algebras, one can in many cases replace  $\text{dexp}^{-1}$  with a low order Lie polynomial while retaining the convergence order of the original Runge–Kutta method. Details of this can be found in [53, 8]. There are also some important cases of Lie algebras for which  $\text{dexp}_\sigma^{-1}$  can be computed exactly in terms of elementary functions, among those is  $\mathfrak{so}(3)$  reported in [15]. Notice that the representation (12) does not depend on the use of the exponential map from  $\mathfrak{g}$  to  $G$ . In principle, one can replace this map with any local diffeomorphism  $\varphi$ , usually one scales  $\varphi$  such that  $\varphi(0) = e$  and  $T_0\varphi = \text{Id}_{\mathfrak{g}}$ . Examples of such maps are the Cayley transformation [23] which can be used for matrix Lie groups of the type  $G_P = \{X \in \mathbf{R}^{d \times d} : X^T P X = P\}$  for a nonsingular  $d \times d$ -matrix  $P$ . These include the orthogonal group  $O(n) = G_I$  and the linear symplectic group  $SP(n) = G_J$  where  $J$  the skew-symmetric matrix of the standard symplectic form. Another possibility is to replace the exponential map by canonical coordinates of the second kind [58].

We present here the well-known Runge–Kutta–Munthe-Kaas method based on the pop-



ular fourth order method of Kutta [36], having Butcher tableau

$$\begin{array}{c|cccc}
 0 & & & & \\
 \frac{1}{2} & \frac{1}{2} & & & \\
 \frac{1}{2} & 0 & \frac{1}{2} & & \\
 1 & 0 & 0 & 1 & \\
 \hline
 & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6}
 \end{array} \tag{16}$$

In the Lie group method, the  $\text{dexp}^{-1}$  map has been replaced by the optimal Lie polynomials.

$$k_1 = hf(y_0),$$

$$k_2 = hf(\exp(\frac{1}{2}k_1) \cdot y_0),$$

$$k_3 = hf(\exp(\frac{1}{2}k_2 - \frac{1}{8}[k_1, k_2]) \cdot y_0),$$

$$k_4 = hf(\exp(k_3) \cdot y_0),$$

$$y_1 = \exp(\frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4 - \frac{1}{2}[k_1, k_4])) \cdot y_0.$$

An important advantage of the Runge–Kutta–Munthe-Kaas schemes is that it is easy to preserve the convergence order when extending them to Lie group integrators. This is not the case with for instance the schemes of Crouch and Grossman [20, 57], where it is necessary to develop order conditions for the nonabelian case. This is also true for the commutator-free methods developed by Celledoni et al. [13]. In fact these methods include those of Crouch and Grossman. The idea here is to allow compositions of exponentials or flows instead of commutator corrections. With stages  $k_1, \dots, k_s$  in the Lie algebra, one includes expressions of the form

$$\exp(\sum_j \beta_j^i k_i) \cdots \exp(\sum_j \beta_j^i k_i) \exp(\sum_j \beta_j^i k_i) y_0$$

both in the definition of the stages and the update itself. In some cases it is also possible to reuse flow calculations from one stage to another, and thereby lower the computational cost of the scheme. An extension of (16) can be obtained as follows, setting  $k_i = hf(Y_i)$

for all  $i$ ,

$$\begin{aligned}
Y_1 &= y_0, \\
Y_2 &= \exp(\tfrac{1}{2}k_1) y_0, \\
Y_3 &= \exp(\tfrac{1}{2}k_2) y_0 \\
Y_4 &= \exp(h(k_3 - \tfrac{1}{2}k_1)) Y_2, \\
y_{\frac{1}{2}} &= \exp(\tfrac{1}{12}(3k_1 + 2k_2 + 2k_3 - k_4)) y_0, \\
y_1 &= \exp(\tfrac{1}{12}(-k_1 + 2k_2 + 2k_3 + 3k_4)) y_{\frac{1}{2}}.
\end{aligned}$$

Note in particular in this example how the expression for  $Y_4$  involves  $Y_2$  and thereby one exponential calculation has been saved.

## 2.2 A plenitude of group actions

We saw in the first examples with Duffing's equation that the manifold  $M$ , the group  $G$  and even the way  $G$  acts on  $M$  can be chosen in different ways. It is not obvious which action is the best or suits the purpose in the problem at hand. Most examples we know from the literature are using matrix Lie groups  $G \subseteq GL(n)$ , but the choice of group action depends on the problem at hand and the objectives of the simulation. We give here several examples of situations where Lie group integrators can be used.

**$G$  acting on  $G$ .** In the case,  $M = G$  it is natural to use either left or right multiplication as the action

$$L_g(m) = g \cdot m \quad \text{or} \quad R_g(m) = m \cdot g, \quad g, m \in G.$$

The correspondence between the vector field  $F \in \mathcal{X}(M)$  and the map (10) is then just the tangent map of left or right multiplication

$$F(g) = T_e L_g(f(g)) \quad \text{or} \quad F(g) = T_e R_g(\tilde{f}(g)), \quad g \in G.$$

When working with matrices, this simply amounts to setting  $F(g) = g \cdot f(g)$  or  $F(g) = \tilde{f}(g) \cdot g$ . Note that  $\tilde{f}(g)$  is related to  $f(g)$  through the adjoint representation of  $G$ ,  $\text{Ad} : G \rightarrow \text{Aut}(\mathfrak{g})$

$$\tilde{f}(g) = \text{Ad}_g f(g), \quad \text{Ad}_g = T_e L_g \circ T_e R_g^{-1}.$$

**The coadjoint action and Lie–Poisson systems.** Lie group integrators for this interesting case were studied by Engø and Faltinsen [25]. Suppose  $G$  is a Lie group and the manifold under consideration is the dual space  $\mathfrak{g}^*$  of its Lie algebra  $\mathfrak{g}$ . The coadjoint action by  $G$  on  $\mathfrak{g}^*$  is denoted  $\text{Ad}_g^*$  defined for any  $g \in G$  as follows:

$$\langle \text{Ad}_g^* \mu, \xi \rangle = \langle \mu, \text{Ad}_g \xi \rangle, \quad \forall \xi \in \mathfrak{g}.$$

It is well known that mechanical systems formulated on the cotangent bundle  $T^*G$  with a left or right invariant Hamiltonian can be reduced to a system on  $\mathfrak{g}^*$  given as

$$\dot{\mu} = \pm \text{ad}_{\frac{\partial H}{\partial \mu}}^* \mu$$

the negative sign is used in case of right invariance. The solution to this system preserves coadjoint orbits, which makes it natural to suggest the group action

$$g \cdot \mu = \text{Ad}_{g^{-1}}^* \mu,$$

so that the resulting Lie group integrator also respects this invariant. For Euler's equations for the free rigid body, the Hamiltonian is left invariant and the coadjoint orbits are spheres in  $\mathfrak{g}^* \cong \mathbf{R}^3$ .

**Homogeneous spaces and the Stiefel and Grassmann manifolds.** The situation when  $G$  acts on itself by left or right multiplication is a special case of a homogeneous space [55], where the assumption is only that  $G$  acts transitively and continuously on some manifold  $M$ . Homogeneous spaces are isomorphic to the quotient  $G/G_x$  where  $G_x$  is the *isotropy group* for the action at an arbitrarily chosen point  $x \in M$

$$G_x = \{y \in G : y \cdot x = x\}.$$

All isotropy groups  $G_x$ ,  $x \in M$  are conjugate and therefore the isomorphism type of  $G/G_x$  is independent of  $x$ . A much encountered example is the hypersphere  $M = S^{d-1}$  corresponding to the left action by  $G = SO(d)$ , the Lie group of orthogonal  $d \times d$  matrices with unit determinant. One has  $S^{d-1} = SO(d)/SO(d-1)$ . We have in fact already discussed the example of the free rigid body (8) where  $M = S^2$ .

The Stiefel manifold  $\text{St}(d, k)$  can be represented by the set of  $d \times k$ -matrices with orthonormal columns. An action on this set is obtained by left multiplication by  $G = SO(d)$ . Lie group integrators for Stiefel manifolds are extensively studied in the literature, see e.g. [16, 35] and some applications involving Stiefel manifolds are discussed in Section 6. An important subclass of the homogeneous spaces is the symmetric spaces, also obtained through a transitive action by a Lie group  $G$ , where  $M = G/G_x$ , but here one requires in addition that the isotropy subgroup is an open subgroup of the fixed point set of an involution of  $G$  [54]. A prominent example of a symmetric space in applications is the Grassmann manifold, obtained as  $SO(d)/(SO(k) \times SO(d-k))$ .

**Isospectral flows.** In isospectral integration one considers dynamical systems evolving on the manifold of  $d \times d$ -matrices sharing the same Jordan form. Considering the case of symmetric matrices, one can use the transitive group action by  $SO(d)$  given as

$$g \cdot m = g^T m g.$$

This action is transitive, since any symmetric matrix can be diagonalized by an appropriately chosen orthogonal matrix. If the eigenvalues are distinct, then the isotropy group is discrete and consists of all matrices in  $SO(d)$  which are diagonal.

Lie group integrators for isospectral flows have been extensively studied, see for example [6, 7]. See also [9] for an application to the KdV equation.

**Tangent and cotangent bundles.** For mechanical systems the natural phase space will often be the tangent bundle  $TM$  as in the Lagrangian framework or the cotangent bundle  $T^*M$  in the Hamiltonian framework. The seminal paper by Lewis and Simo [42] discusses several Lie group integrators for mechanical systems on cotangent bundles, deriving methods which are symplectic, energy and momentum preserving. Engø [24] suggested a way to generalize the Runge–Kutta–Munthe-Kaas methods into a partitioned version when  $M$  is a Lie group. Marsden and collaborators have developed the theory of Lie group integrators from the variational viewpoint over the last two decades, for an overview, see [45]. For more recent work pertaining to Lie groups in particular, see [38, 3, 61]. In Section 3 we present what we believe to be the first symplectic partitioned Lie group integrators on  $T^*G$  phrased in the framework we have discussed here. Considering trivialized cotangent bundles over Lie groups is particularly attractive since there is a natural way to extend action by left multiplication from  $G$  to  $G \times \mathfrak{g}^*$  via (19).

**The affine group and its use in semi-linear PDE methods.** Lie group integrators can also be used for approximating the solution to partial differential equations, the most obvious choice of PDE model being the semilinear problem

$$u_t = Lu + N(u), \tag{17}$$

where  $L$  is a linear differential operator and  $N(u)$  is some nonlinear map, typically containing derivatives of lower order than  $L$ . After discretizing in space, (17) is turned into a system of  $n_d$  ODEs, for some large  $n_d$ ,  $L$  becomes an  $n_d \times n_d$ -matrix, and  $N : \mathbf{R}^{n_d} \rightarrow \mathbf{R}^{n_d}$  a nonlinear function. We may now as in [52] introduce the action on  $\mathbf{R}^{n_d}$  by some subgroup of the affine group represented as the semidirect product  $G = GL(n_d) \ltimes \mathbf{R}^{n_d}$ . The group product, identity, and inverse are given as

$$(A_1, b_1) \cdot (A_2, b_2) = (A_1 \cdot A_2, A_1 b_2 + b_1), \quad e = (I, 0), \quad (A, b)^{-1} = (A^{-1}, -A^{-1}b).$$

The action on  $\mathbf{R}^{n_d}$  is

$$(A, b) \cdot x = Ax + b, \quad (A, b) \in G, \quad x \in \mathbf{R}^{n_d}$$

and the Lie algebra and commutator are given as

$$\mathfrak{g} = (\xi, c), \quad \xi \in \mathfrak{gl}(n_d), \quad c \in \mathbf{R}^{n_d}, \quad [(\xi_1, c_1), (\xi_2, c_2)] = ([\xi_1, \xi_2], \xi_1 c_2 - \xi_2 c_1 + c_1).$$

In many interesting PDEs, the operator  $L$  is constant, so it makes sense to consider the  $n_d + 1$ -dimensional subalgebra  $\mathfrak{g}_L$  of  $\mathfrak{g}$  consisting of elements  $(\alpha L, c)$  where  $\alpha \in \mathbf{R}$ ,  $c \in \mathbf{R}^d$ , so that the map  $f : \mathbf{R}^{n_d} \rightarrow \mathfrak{g}_L$  is given as

$$f(u) = (L, N(u))$$

One parameter subgroups are obtained through the exponential map as follows

$$\exp(t(L, b)) = (\exp(tL), \phi(tL)tb).$$

Here the entire function  $\phi(z) = (\exp(z) - 1)/z$  familiar from the theory of exponential integrators appears. As an example, one could now consider the Lie–Euler method in this setting, which coincides with the exponential Euler method

$$u_1 = \exp(h(L, N(u_0))) \cdot u_0 = \exp(hL)u_0 + h\phi(hL)N(u_0)$$

There is a large body of literature on exponential integrators, going almost half a century back in time, see [32] and the references therein for an extensive account.

### 2.3 Isotropy – challenges and opportunities

An issue which we have already mentioned a few times is that the map  $\lambda_* : \mathfrak{g} \rightarrow \mathcal{X}(M)$  defined in (9) is not necessarily injective. This means that the choice of  $f : M \rightarrow \mathfrak{g}$  is not unique. In fact, if  $g : M \rightarrow \mathfrak{g}$  is any map satisfying  $\lambda_*(g(m))(m) = 0$  for all  $m \in M$ , then we could replace the map  $f$  by  $f + g$  in (10) without altering the vector field  $F$ . But such a modification of  $f$  *will have* an impact on the numerical schemes that we consider. This freedom in the setup of the methods makes it challenging to prove general results for Lie group methods, it might seem that some restrictions should apply to the isotropy choice for a more well defined class of schemes. However, the freedom can of course also be taken advantage of to obtain approximations of improved quality.

An illustrative example is the two-sphere  $S^2$  acted upon linearly by the special orthogonal group  $SO(3)$ . Representing elements of the Lie algebra  $\mathfrak{so}(3)$  by vectors in  $\mathbf{R}^3$ , and points on the sphere as unit length vectors in  $\mathbf{R}^3$ , we may facilitate (10) as

$$F(m) = f(m) \times m = (f(m) + \alpha(m)m) \times m,$$

for any scalar function  $\alpha : S^2 \rightarrow \mathbf{R}$ . Using for instance the Lie–Euler method one would get

$$m_1 = \exp(f(m_0) + \alpha(m_0)m_0) \cdot m_0, \tag{18}$$

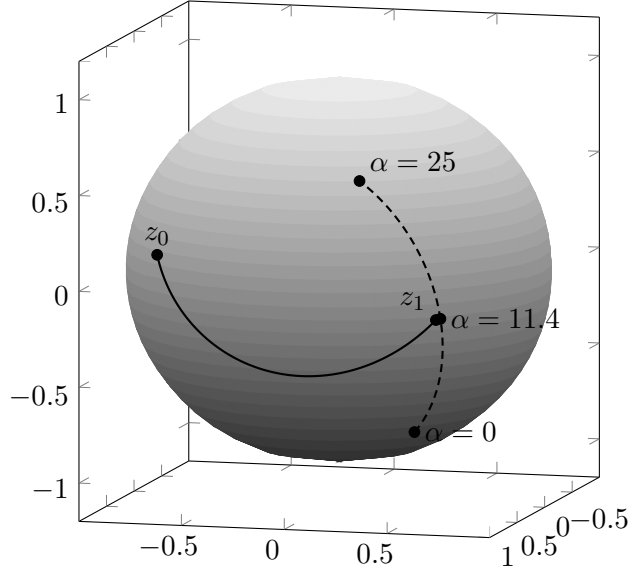


Figure 2: The effect of isotropy on  $S^2$  for Euler’s free rigid body equations.

where the  $\exp$  is the matrix exponential of the  $3 \times 3$  skew-symmetric matrix associated to a vector in  $\mathbf{R}^3$  via the “hat map”. Clearly the approximation depends on the choice of  $\alpha(m)$ . The approach of Lewis and Olver [41] was to use the isotropy to improve certain qualitative features of the solution. In particular, they studied how the orbital error could be reduced by choosing the isotropy in a clever way. In Figure 2 we illustrate the issue of isotropy for the Euler free rigid body equations. The curve drawn from the initial point  $z_0$  to  $z_1$  is the exact solution, i.e. the momenta in body coordinates. The broken line shows the terminal points using the Lie–Euler method for  $\alpha$  varying between 0 and 25.

Another potential difficulty with isotropy is the increased computational complexity when the group  $G$  has much higher dimension than the manifold  $M$ . This could for instance be the case with the Stiefel manifold  $\text{St}(d, k)$  if  $d \gg k$ . Linear algebra operations used in integrating differential equations on Stiefel manifold should preferably be of complexity  $\mathcal{O}(dk^2)$ . But solving a corresponding problem in the Lie algebra  $\mathfrak{so}(d)$  would typically require linear algebra operations of complexity  $\mathcal{O}(d^3)$ . By taking advantage of the many degrees of freedom provided by the isotropy, it is actually possible to reduce the cost down to the required  $\mathcal{O}(dk^2)$  operations as explained for instance in [15] and [35].

### 3 Symplectic integrators on the cotangent bundle of a Lie group

In this section we shall assume that the manifold is the cotangent bundle  $T^*G$  of a Lie group  $G$ . Let  $R_g: G \rightarrow G$  be the right multiplication operator such that  $R_g(h) = hg$  for any  $h \in G$ . The tangent map of  $R_g$  is denoted  $R_{g*} := TR_g$ . Any cotangent vector  $p_g \in T_g^*G$  can be associated to  $\mu \in \mathfrak{g}^*$  by right trivialization as follows: Write  $v_g \in T_gG$  in the form  $v_g = R_{g*}\xi$  where  $\xi \in \mathfrak{g}$ , so that  $\langle p_g, v_g \rangle = \langle p_g, R_{g*}\xi \rangle = \langle R_g^*p_g, \xi \rangle$ , where we have used  $R_g^*$  for the dual map of  $R_{g*}$ . We therefore represent  $p_g \in T_g^*G$  by  $\mu = R_g^*p_g \in \mathfrak{g}^*$ . Thus, we may use as phase space  $G \times \mathfrak{g}^*$  rather than  $T^*G$ . For applying Lie group integrators we need a transitive group action on  $G \times \mathfrak{g}^*$  and this can be achieved by lifting the group structure of  $G$  and use left multiplication in the extended group. The semi-direct product structure on  $\mathbf{G} := G \ltimes \mathfrak{g}^*$  is defined as

$$(g_1, \mu_1) \cdot (g_2, \mu_2) = (g_1 \cdot g_2, \mu_1 + \text{Ad}_{g_1}^* \mu_2) \quad (19)$$

Similarly, the tangent map of right multiplication extends as

$$TR_{(g,\mu)}(R_{h*}\zeta, \nu) = (R_{hg*}\zeta, \nu - \text{ad}_\zeta^* \text{Ad}_{h^{-1}}^* \mu), \quad g, h \in G, \zeta \in \mathfrak{g}, \mu, \nu \in \mathfrak{g}^*.$$

Of particular interest is the restriction of  $TR_{(g,\mu)}$  to  $T_e\mathbf{G} \cong \mathfrak{g} \ltimes \mathfrak{g}^*$ .

$$T_eR_{(g,\mu)}(\zeta, \nu) = (R_{g*}\zeta, \nu - \text{ad}_\zeta^* \mu)$$

The natural symplectic form on  $T^*G$  is defined as

$$\Omega_{(g,p_g)}((\delta v_1, \delta \pi_1), (\delta v_2, \delta \pi_2)) = \langle \delta \pi_2, \delta v_1 \rangle - \langle \delta \pi_1, \delta v_2 \rangle,$$

and by right trivialization it may be pulled back to  $G \ltimes \mathfrak{g}^*$  and then takes the form

$$\omega_{(g,\mu)}((R_{g*}\xi_1, \delta \nu_1), (R_{g*}\xi_2, \delta \nu_2)) = \langle \delta \nu_2, \xi_1 \rangle - \langle \delta \nu_1, \xi_2 \rangle - \langle \mu, [\xi_1, \xi_2] \rangle \quad (20)$$

The presentation of differential equations on  $T^*G$  is now achieved via the action by left multiplication, meaning that any vector field  $F \in \mathcal{X}(G \ltimes \mathfrak{g}^*)$  is expressed by means of a map  $f: G \ltimes \mathfrak{g}^* \rightarrow \mathfrak{g} \ltimes \mathfrak{g}^*$

$$F(g, \mu) = T_eR_{(g,\mu)}f(g, \mu) = (R_{g*}f_1, f_2 - \text{ad}_{f_1}^* \mu) \quad (21)$$

where  $f_1 = f_1(g, \mu) \in \mathfrak{g}$ ,  $f_2 = f_2(g, \mu) \in \mathfrak{g}^*$  are the two components of  $f$ . We are particularly interested in the case that  $F$  is a Hamiltonian vector field which means that  $F$  satisfies the relation

$$\mathbf{i}_F \omega = dH \quad (22)$$

for some Hamiltonian function  $H: T^*G \rightarrow \mathbf{R}$ . A simple calculation using (20), (21) and (22) shows that the corresponding map  $f$  for such a Hamiltonian vector field is

$$f(g, \mu) = \left( \frac{\partial H}{\partial \mu}(g, \mu), -R_g^* \frac{\partial H}{\partial g}(g, \mu) \right)$$

We have come up with the following family of symplectic Lie group integrators on  $G \times \mathfrak{g}^*$

$$(\xi_i, \bar{n}_i) = hf(G_i, M_i), \quad n_i = \text{Ad}_{\exp(X_i)}^* \bar{n}_i, \quad i = 1, \dots, s.$$

$$(g_1, \mu_1) = \exp(Y, \text{dexp}_Y^{-1*} \sum_{i=1}^s b_i n_i) \cdot (g_0, \mu_0)$$

$$Y = \sum_{i=1}^s b_i \xi_i, \quad X_i = \sum_{j=1}^s a_{ij} \xi_j, \quad i = 1, \dots, s.$$

$$G_i = \exp(X_i) \cdot g_0, \quad i = 1, \dots, s$$

$$M_i = \text{dexp}_{-Y}^*(\mu) + \sum_{j=1}^s (b_j \text{dexp}_{-Y}^* - \frac{b_j a_{ji}}{b_i} \text{dexp}_{-X_j}^*) n_j, \quad i = 1, \dots, s.$$

It is assumed that  $\sum_{i=1}^s b_i = 1$  and that  $b_i \neq 0$ ,  $1 \leq i \leq s$ . The symplecticity of these schemes is a consequence of their derivation from a variational principle, following ideas similar to that of [3] and [61]. One should be aware that order barriers for this type of schemes may apply, and that further stage corrections may be necessary to obtain high order methods. In Figure 3 we show numerical experiments for the heavy top where the Hamiltonian is given as

$$H(g, \mu) = \frac{1}{2} \langle \mu, \mathbb{I}^{-1} \mu \rangle + e_3^T g u_0,$$

where  $\mathbb{I}: \mathfrak{g} \rightarrow \mathfrak{g}^*$  is the inertia tensor, here represented as a diagonal  $3 \times 3$  matrix,  $u_0$  is the initial position of the top's center of mass, and  $e_3$  is the canonical unit vector in the vertical direction. We have chosen  $\mathbb{I} = 10^3 \text{diag}(1, 5, 6)$  and  $u_0 = e_3$ . The initial values used were  $g_0 = I$  (identity matrix) and  $\mu_0 = 10 \mathbb{I}(1, 1, 1)^T$ . We have used the coefficients of the well known  $\theta$ -method, where  $s = 1$ ,  $b_1 = 1$ ,  $a_{11} = \theta$ , and we compare the behaviour of the symplectic schemes presented here to the Runge–Kutta–Munthe–Kaas (RKMK) method with the same  $\theta$ -value. In Figure 3 we have drawn the time evolution of the center of mass,  $u_n = g_n \cdot u_0$ , the characteristic band structure observed for the symplectic methods was reported in [17]. The RKMK method with  $\theta = \frac{1}{2}$  exhibits a similar behaviour, but the bands are expanding faster than for the symplectic ones. We have also found in these experiments that none of the symplectic schemes,  $\theta = 0$  and  $\theta = \frac{1}{2}$  have energy drift, but this is also the case for the RKMK method with  $\theta = \frac{1}{2}$ . For  $\theta = 0$ , however, the RKMK



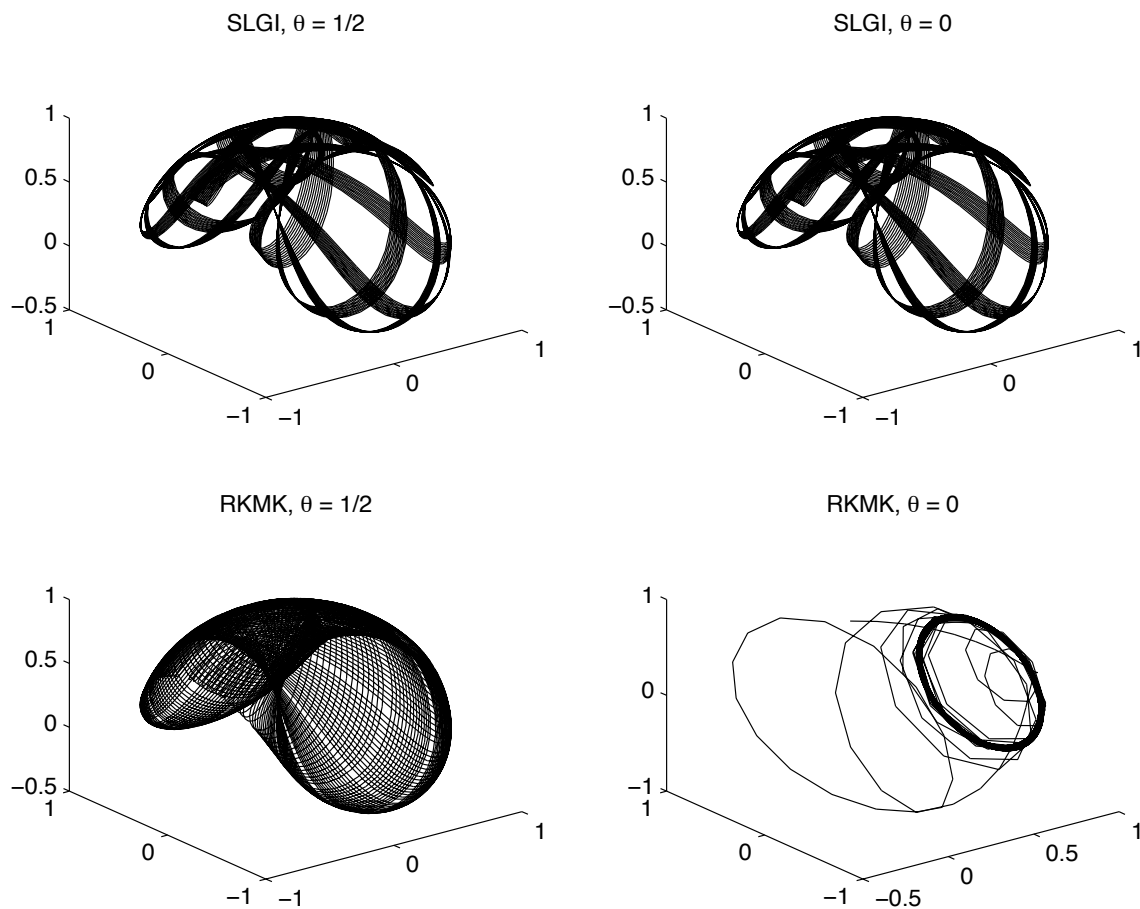


Figure 3: Heavy top simulations with the symplectic (SLGI) methods and RKMK  $\theta$ -methods with  $\theta = 0, \frac{1}{2}$ .

method shows energy drift as expected. These tests were done with step size  $h = 0.05$  over  $10^5$  steps.

## 4 Discrete gradients and integral preserving methods on Lie groups

The discrete gradient method for preserving first integrals has to a large extent been made popular through the works of Gonzalez [27] and McLachlan et al. [46], the latter proved the result that under relatively general circumstances, a differential equation which has a first integral  $I(x)$  can be written in the form

$$\dot{x} = S(x)\nabla I(x)$$

for some non-unique solution-dependent skew-symmetric matrix  $S(x)$ . The idea is to introduce a mapping which resembles the true gradient; a *discrete gradient*  $\bar{\nabla}I: \mathbf{R}^d \times \mathbf{R}^d \rightarrow \mathbf{R}^d$  is a continuous map which satisfies the following two conditions

$$\bar{\nabla}I(x, x) = \nabla I(x), \quad \forall x,$$

$$I(y) - I(x) = \bar{\nabla}^T I(x, y)(y - x), \quad \forall x \neq y.$$

An integrator which preserves  $I$ , that is,  $I(x_n) = I(x_0)$  for all  $n$  is now easily devised as

$$\frac{x_{n+1} - x_n}{h} = \tilde{S}(x_n, x_{n+1})\bar{\nabla}I(x_n, x_{n+1})$$

where  $\tilde{S}(x, y)$  is some consistent approximation to  $S(x)$ , i.e.  $\tilde{S}(x, x) = S(x)$ . There exist several discrete gradients, two of the most popular are

$$\bar{\nabla}I(x, y) = \int_0^1 \nabla I(\zeta y + (1 - \zeta)x) d\zeta \quad (23)$$

and

$$\bar{\nabla}I(x, y) = \nabla I\left(\frac{x+y}{2}\right) + \frac{I(y) - I(x) - \nabla I\left(\frac{x+y}{2}\right)^T(y-x)}{\|y-x\|^2}(y-x). \quad (24)$$

The matrix  $\tilde{S}(x, y)$  can be constructed with the purpose of increasing the accuracy of the resulting approximation, see e.g. [59].

We now generalize the concept of the discrete gradient to a Lie group  $G$ . We consider differential equations which can, for a given dual two-form<sup>1</sup>  $\omega \in \Omega_2(G)$  and a function  $H: G \rightarrow \mathbf{R}$  be written in the form

$$\dot{x} = \mathbf{i}_{dH}\omega \quad (25)$$

---

<sup>1</sup>By dual two-form, we here mean a differential two-form on  $G$  such that on each fiber of the cotangent bundle we have  $\omega|_x: T_x^*G \times T_x^*G \rightarrow \mathbf{R}$ , a bilinear, skew-symmetric form.

where  $\mathbf{i}_\alpha$  is the interior product  $\mathbf{i}_\alpha\omega(\beta) = \omega(\alpha, \beta)$  for any two one-forms  $\alpha, \beta \in \Omega^1(G)$ . The function  $H$  is a first integral since

$$\frac{d}{dt}H(x(t)) = dH|_{x(t)}(\dot{x}(t)) = \omega(dH, dH) = 0.$$

We define the *trivialized discrete differential* (TDD) of the function  $H$  to be a continuous map  $\bar{d}H : G \times G \rightarrow \mathfrak{g}^*$  such that

$$H(x') - H(x) = \langle \bar{d}H(x, x'), \log(x'x^{-1}) \rangle$$

$$\bar{d}H(x, x) = R_x^* dH|_x$$

A numerical method can now be defined in terms of the discrete differential as

$$x' = \exp(h \mathbf{i}_{\bar{d}H(x, x')} \bar{\omega}(x, x')) x$$

where  $\bar{\omega} : G \times G \rightarrow \Lambda^2(\mathfrak{g}^*)$  is continuous. This exterior form is some local trivialized approximation to  $\omega$ . For consistency we require

$$\bar{\omega}(x, x)(R_x^* \alpha, R_x^* \beta) = \omega|_x(\alpha, \beta), \quad \alpha, \beta \in T_x^* G. \quad (26)$$

We easily see that this method preserves  $H$  exactly, since

$$\begin{aligned} H(x') - H(x) &= \langle \bar{d}H(x, x'), \log(x'x^{-1}) \rangle \\ &= \langle \bar{d}H(x, x'), h \mathbf{i}_{\bar{d}H(x, x')} \bar{\omega} \rangle \\ &= h \bar{\omega}(\bar{d}H(x, x'), \bar{d}H(x, x')) = 0. \end{aligned}$$

Extending (23) to the Lie group setting we define the following TDD

$$\bar{d}H(x, x') = \int_0^1 R_{\ell(\xi)}^* dH|_{\ell(\xi)} d\xi, \quad \ell(\xi) = \exp(\xi \log(x' \cdot x^{-1}))x. \quad (27)$$

Similarly, for any given inner product on  $\mathfrak{g}$ , we may extend the discrete gradient (24) to

$$\bar{d}H(x, x') = R_{\bar{x}}^* dH|_{\bar{x}} + \frac{H(x') - H(x) - \langle R_{\bar{x}}^* dH|_{\bar{x}}, \eta \rangle}{\|\eta\|^2} \eta^\flat, \quad \eta = \log(x'x^{-1}), \quad (28)$$

where  $\bar{x} \in G$  could for instance be  $\bar{x} = \exp(\eta/2)x$ , a choice which would cause  $\bar{d}H(x, x') = \bar{d}H(x', x)$ .

Suppose that the ODE vector field  $F$  is known as well as the invariant  $H$ . A dual two-form  $\omega$  can now be defined in terms of a Riemannian metric on  $G$ . By index raising applied to  $dH$ , we obtain the Riemannian gradient vector field  $\text{grad } H$ , and we define

$$\omega = \frac{\text{grad } H \wedge F}{\|\text{grad } H\|^2} \quad \Rightarrow \quad \mathbf{i}_{dH}\omega = F.$$

**Example:** We consider the equations for the attitude rotation of a free rigid body expressed using Euler parameters. The set  $S^3 = \{\mathbf{q} \in \mathbf{R}^4 \mid \|\mathbf{q}\|_2 = 1\}$  with  $\mathbf{q} = [q_0, \mathbf{q}]$  ( $q_0 \in \mathbf{R}$  and  $\mathbf{q} \in \mathbf{R}^3$ ), is a Lie group with the quaternion product

$$\mathbf{p} \cdot \mathbf{q} = [p_0 q_0 - \mathbf{p}^T \mathbf{q}, p_0 \mathbf{p} + q_0 \mathbf{q} + \mathbf{p} \times \mathbf{q}]$$

with unit  $\mathbf{e} = [1 \ 0 \ 0 \ 0]$  and inverse  $\mathbf{q}_c = [q_0, -\mathbf{q}]$ . Denote by “ $\hat{\cdot}$ ” the hat-map identifying  $\mathbf{R}^3$  with  $\mathfrak{so}(3)$ :

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \mapsto \hat{\mathbf{v}} = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}.$$

The Lie group  $S^3$  can be mapped into  $SO(3)$  by the Euler–Rodrigues map:

$$\mathcal{E}(\mathbf{q}) = I_3 + 2q_0 \hat{\mathbf{q}} + 2\hat{\mathbf{q}}^2,$$

where  $I_3$  denotes the  $3 \times 3$  identity matrix. The Lie algebra  $\mathfrak{s}^3$  of  $S^3$  is the set of so called pure quaternions, the elements of  $\mathbf{R}^4$  with first component equal to zero, identifiable with  $\mathbf{R}^3$  and with  $\mathfrak{so}(3)$  via the hat-map.

The equations of the attitude rotation of a free rigid body on  $S^3$  read

$$\dot{\mathbf{q}} = f(\mathbf{q}) \cdot \mathbf{q}, \quad f(\mathbf{q}) = \mathbf{q} \cdot \mathbf{v} \cdot \mathbf{q}_c,$$

and

$$\mathbf{v} = [0, \mathbf{v}], \quad \mathbf{v} = \frac{1}{2} \mathbb{I}^{-1} \mathcal{E}(\mathbf{q}_c) \mathbf{m}_0,$$

$\mathbf{m}_0$  is the initial body angular momentum and  $\mathbb{I}$  is the diagonal inertia tensor, and according with the notation previously used in this section  $F(\mathbf{q}) = f(\mathbf{q}) \cdot \mathbf{q}$ . The energy function is

$$H(\mathbf{q}) = \frac{1}{2} \mathbf{m}_0^T \mathcal{E}(\mathbf{q}) \mathbb{I}^{-1} \mathcal{E}(\mathbf{q}_c) \mathbf{m}_0.$$

We consider the  $\mathbf{R}^3$  Euclidean inner product as metric in the Lie algebra  $\mathfrak{s}^3$ , and obtain by right translation a Riemannian metric on  $S^3$ . The Riemannian gradient of  $H$  with respect to this metric is then

$$\text{grad } H = (I_4 - \mathbf{q} \mathbf{q}^T) \nabla H,$$

where  $I_4$  is the identity in  $\mathbf{R}^4$  and  $\nabla H$  is the usual gradient of  $H$  as a function from  $\mathbf{R}^4$  to  $\mathbf{R}$ . We identify  $\mathfrak{s}^3$  with its dual, and using  $\text{grad } H$  in (24) we obtain the (dual) discrete differential  $\overline{\text{grad } H}(\mathbf{q}, \mathbf{q}') \in \mathfrak{s}^3$ .

The two-form  $\omega = \frac{\text{grad } H \wedge F}{\|\text{grad } H\|^2}$  with respect to the right trivialization can be identified with the  $4 \times 4$  skew-symmetric matrix

$$\omega_R(\mathbf{q}) = \frac{\xi \gamma^T - \gamma \xi^T}{\|\gamma\|^2}, \quad \xi, \gamma \in \mathfrak{s}^3, \quad \xi \cdot \mathbf{q} = F(\mathbf{q}), \quad \gamma \cdot \mathbf{q} = \text{grad } H(\mathbf{q}),$$

$\omega_R(\mathbf{q})$  has first row and first column equal to zero. We choose  $\bar{\omega}$  to be

$$\bar{\omega}(\mathbf{q}, \mathbf{q}') = \omega_R(\bar{\mathbf{q}}), \quad \bar{\mathbf{q}} = \exp(\eta/2)\mathbf{q}, \quad \eta = \log(\mathbf{q}' \cdot \mathbf{q}_c),$$

i.e.  $\omega_R$  frozen at the mid-point  $\bar{\mathbf{q}}$ . The energy-preserving Lie group method of second order is

$$\mathbf{q}' = \exp(h \bar{\omega}(\mathbf{q}, \mathbf{q}') \overline{\text{grad } H}(\mathbf{q}, \mathbf{q}')) \cdot \mathbf{q},$$

and  $\exp$  is the exponential map from  $\mathfrak{s}^3$  to  $S^3$ ,  $\log: S^3 \rightarrow \mathfrak{s}^3$  is its inverse.

In Figure 4 we plot the body angular momentum vector  $\mathbf{m} = \mathcal{E}(\mathbf{q}_c)\mathbf{m}_0$  on a time interval  $[0, T]$ ,  $T = 1000$ , for four different methods: the Lie group energy-preserving integrator just described (top left), the built-in MATLAB routine `ode45` with absolute and relative tolerance  $10^{-6}$  (top right); the `ode45` routine with tolerances  $10^{-14}$  (bottom left); and the explicit Heun RKMK Lie group method (bottom right). The two Lie group methods both have order 2. The energy preserving method is both symmetric, energy preserving and it preserves the constraint  $\|\mathbf{q}\|_2 = 1$ . The Lie group integrators use a step-size  $h = 1/64$ . The solution of the built-in MATLAB routine at high precision is qualitatively similar to the highly accurate solution produced by MATLAB with tolerances  $10^{-14}$ , the energy error is also comparable for these two experiments. The performance of other MATLAB built-in routines we tried was worse than for `ode45`. We remark that the equations are formulated as differential equations on  $S^3$ , a formulation of the problem in form of a differential algebraic equation would possibly have improved the performance of the MATLAB built-in routines. However it seems that the preservation of the constraint alone can not guarantee the good performance of the method. In fact the explicit (non-symmetric) Lie group integrator preserves the constraint  $\|\mathbf{q}\|_2 = 1$ , but performs poorly on this problem (see Figure 4 bottom right). The cost per step of the explicit Lie group integrator is much lower than for the energy-preserving symmetric Lie group integrator.

## 5 Applications to nonlinear problems of evolution in classical mechanics

The emphasis on the use of Lie groups in modeling and simulation of engineering problems in classical mechanics started in the eighties with the pioneering and fundamental work of

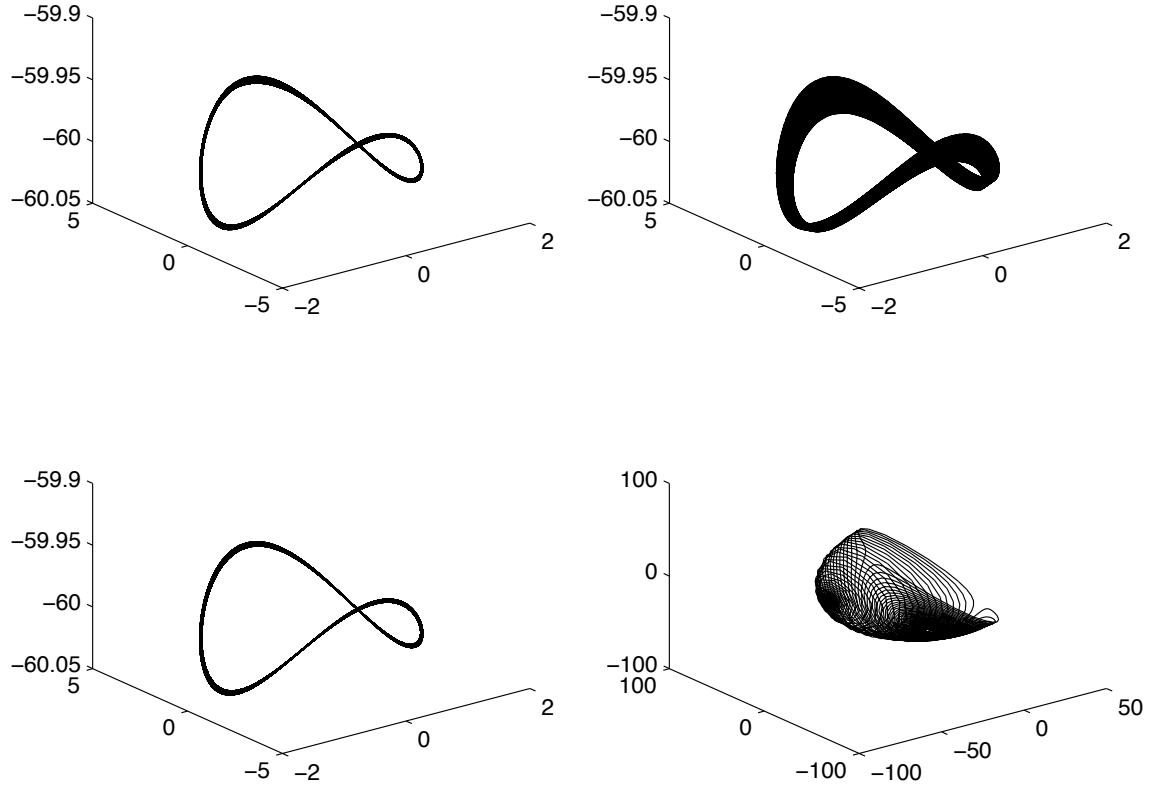


Figure 4: Free rigid body angular momentum, time interval  $[0, 1000]$ , moments of inertia  $I_1 = 1$ ,  $I_2 = 5$ ,  $I_3 = 60$ , initial angular velocity  $\mathbb{I} \mathbf{m}_0 = [1, 1/2, -1]$ : (top left) energy-preserving Lie group method,  $h = 1/64$ ; (top right) `ode45` with tolerances  $10^{-6}$ ; (bottom left) `ode45` with tolerances  $10^{-14}$ ; (bottom right) Heun RKMK,  $h = 1/64$ .

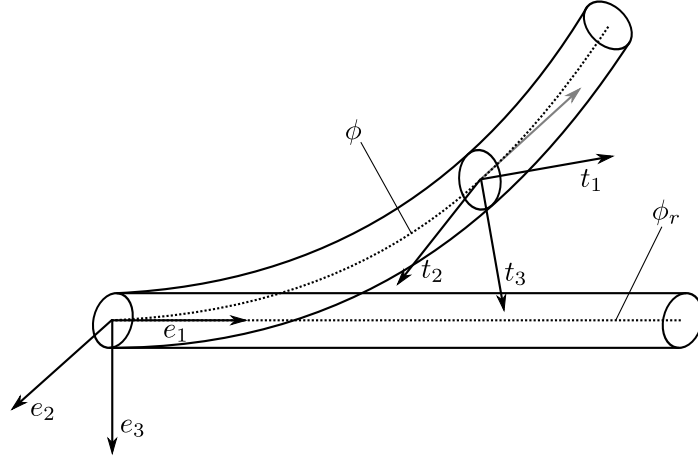


Figure 5: Geometric rod model.  $\phi$  line of centroids, cross section identified by the frame  $[\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3]$ .

J.C. Simo and his collaborators. In the case of rod dynamics, for example, models based on partial differential equations were considered where the configuration of the centreline of the rod is parametrized via arc-length, and the movement of a rigid frame attached to each of the cross sections of the rod is considered, see Figure 5. This was first presented in a geometric context in [64].

In robot technology, especially robot locomotion and robot grasping, the occurrence of non-holonomically constrained models is very common. The motion of robots equipped with wheels is not always locally controllable, but is often globally controllable. A classical example is the parking of a car that cannot be moved in the direction orthogonal to its wheels. The introduction of Lie groups and Lie brackets to describe the dynamics of such systems, has been considered by various authors, see for example [56]. The design of numerical integration methods in this context has been addressed in the paper of Crouch and Grossman, [20]. These methods have had a fundamental impact to the successive developments in the field of Lie group methods.

The need for improved understanding of non-holonomic numerical integration has been for example advocated in [47]. Recent work in this field has led to the construction of low order non-holonomic integrators based on a discrete Lagrange-d'Alembert principle, [19, 49]. The use of Lie group integrators in this context has been considered in [40, 49].

We have already mentioned the relevance of rigid body dynamics to the numerical discretization of rod models. There are many other research areas in which the accurate and efficient simulation of rigid body dynamics is crucial: molecular dynamics, satellite dynamics, and celestial mechanics just to name a few, [39]. In some of these applications, it is desirable to produce numerical approximations which are accurate possibly to the size of roundoff. The simulations of interest occur over very long times and/or a large

number of bodies and this inevitably causes propagation of errors even when the integrator is designed to be very accurate. For this reason accurate symplectic rigid body integrators are of interest because they can guarantee that the roundoff error produced by the accurate computations can stay bounded also in long time integration. This fact seems to be of crucial importance in celestial mechanics, [37]. A symplectic and energy preserving Lie group integrator for the free rigid body motion was proposed in [42]. The method computes a time re-parametrization of the exact solution. Some recent and promising work in this field has been presented in [48, 17, 10, 29]. The control of rigid bodies with variational Lie group integrators was considered in [40].

In the next section we illustrate the use of Lie group methods in applications on a particular case study, the pipe-laying process from ships to the bottom of the sea.

## 5.1 Rigid body and rod dynamics

### 5.1.1 Pipe-laying problem

The simulation of deep-water risers, pipelines and drill rigs requires the use of models of long and thin beams subject to large external forces. These are complicated nonlinear systems with highly oscillatory components. We are particularly interested in the correct and accurate simulation of the pipe-laying process from ships on the bottom of the sea, see Figure 6. The problem comprises the modeling of two interacting structures: a long and thin pipe (modeled as a rod) and a vessel (modeled as a rigid body). The system is subject to environmental forces (such as sea and wind effects). The control parameters for this problem are the vessel position and velocity, the pay-out speed and the pipe tension while the control objectives consist in determining the touchdown position of the pipe as well as ensuring the integrity of the pipe and to avoid critical deformations, [34, 62].

The vessel rigid body equations are determining the boundary conditions of the rod. They can be integrated numerically with a splitting and composition technique where the vessel equations are split into a free rigid body part and a damping and control part. The free rigid body equations can be solved with a method proposed in [17] where the angular momentum is accurately and efficiently computed by using Jacobi elliptic functions, the attitude rotation is obtained using a Lie group method based on the Magnus expansion, and the control and damping part are solved exactly.

Simulations of the whole pipe-lay problem with local parametrizations of the pipe and the vessel based on Euler angles have been obtained in [34].

### 5.1.2 Rod dynamics

At fixed time each cross section of the pipe is the result of a rigid rotation in space of a reference cross section, and, analogously, for each fixed value of the space variable the corresponding cross section evolves in time as a forced rigid body, see Figure 5.



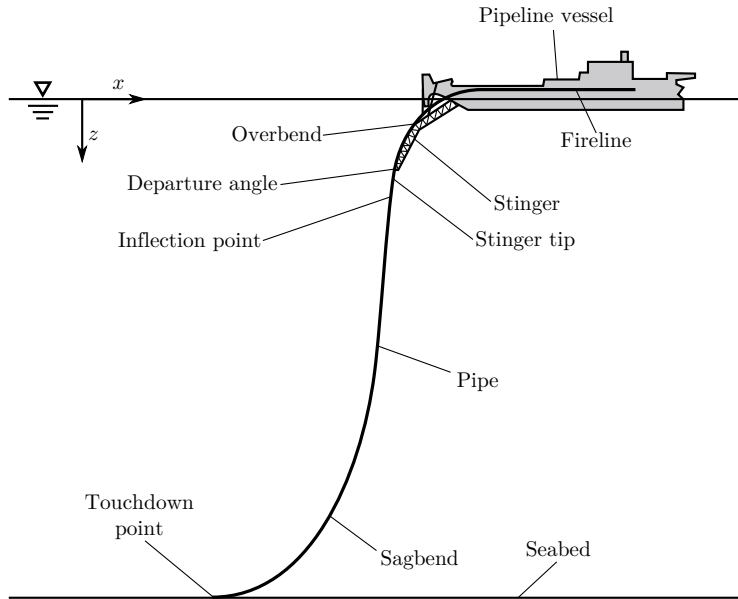


Figure 6: The pipe-lay process.

In [66], partitioned Newmark integrators, of Lie group type, and of moderate order were considered for this problem. While classical Newmark methods are variational integrators and as such are symplectic when applied to Hamiltonian systems [45], the particular approach of [66] involves the use of exponentials for the parametrization of the Lie group  $SO(3)$  and, the symplecticity in this case is not easily achieved. Moreover since the model is a partial differential equation space and time discretizations should be designed so that the overall discrete equations admit solutions and are stable. It turns out that conventional methods perform poorly on such problem in long time simulations. To obtain stable methods reliable in long-time simulation, an energy-momentum method was proposed for the rod problem in [65]. Later this line of thought has been further developed in [60]. The Hamiltonian formulation of this model allows to derive natural structure preserving discretizations into systems of coupled rigid bodies [43].

Following the geometric space-time integration procedure proposed in [26], a multi-Hamiltonian formulation<sup>2</sup> of these equations has been proposed in [18], using the Lie group of Euler parameters. The design of corresponding multi-symplectic Lie group discretizations is still under investigation.

<sup>2</sup>For a definition of the multi-symplectic structure of Hamiltonian partial differential equations, see [4].

## 6 Applications to problems of data analysis and statistical signal processing

The solution of the many-body Schrödinger eigenvalue problem,

$$\hat{\mathbf{H}}\Psi = E\Psi, \quad (29)$$

where the so called electronic ground state (the smallest eigenstate) is sought, is an important problem of computational chemistry. The main difficulty is the curse of dimensionality. Since  $\hat{\mathbf{H}}$  is a differential operator in several space dimensions, a realistic simulation of (29) would require the numerical discretization and solution of a partial differential equation in several space dimensions. The number of space dimensions grows with the number of electrons included in the simulation.

The eigenvalue problem admits an alternative variational formulation. Instead of looking for the smallest eigenvalue and eigenfunction of the Schrödinger equation, one minimizes directly the ground state energy.

After appropriate spatial discretization, the problem becomes a minimization problem on a Riemannian manifold  $\mathcal{M}$ ,

$$\min_{x \in \mathcal{M}} \phi(x), \quad (30)$$

where  $\phi: \mathcal{M} \rightarrow \mathbf{R}$  is a smooth discrete energy function to be minimized on  $\mathcal{M}$ . The discrete energy  $\phi$  considered here is the so called Kohn–Sham energy, [1]. For a related application of Lie group techniques in quantum control see [21].

The general optimization problem giving rise to (30) arises in several applied fields ranging from engineering to applied physics and medicine. Some specific examples are principal component/subspace analysis, eigenvalue and generalized eigenvalue problems, optimal linear compression, noise reduction, signal representation and blind source separation.

### 6.1 Gradient-based optimization on Riemannian manifolds

Let  $\mathcal{M}$  be a Riemannian manifold with metric  $\langle \cdot, \cdot \rangle_{\mathcal{M}}$  and  $\phi: \mathcal{M} \rightarrow \mathbf{R}$  be a smooth cost function to be minimized on  $\mathcal{M}$ , and we want to solve (30). The optimization method based on gradient flow, (written for the minimization problem only, for the sake of easy reading), consists in setting up the differential equation on the manifold,

$$\dot{x}(t) = -\text{grad } \phi(x(t)), \quad (31)$$

with appropriate initial condition  $x(0) = x_0 \in \mathcal{M}$ . The equilibria of equation (31) are the critical points of the function  $\phi$ . In the above equation, the symbol  $\text{grad } \phi$  denotes the Riemannian gradient of the function  $\phi$  with respect to the chosen metric. Namely,  $\text{grad } \phi(x) \in T_x \mathcal{M}$  and  $\phi'|_x(v) = \langle \text{grad } \phi(x), v \rangle_{\mathcal{M}}$  for all  $v \in T_x \mathcal{M}$ .

The solution of (31) on  $\mathcal{M}$  may be locally expressed in terms of a curve on the tangent space  $T_{x_0}\mathcal{M}$  using a retraction map  $\varphi$ . Retractions are tangent space parametrizations of  $\mathcal{M}$ , and allow us to write

$$x(t) = \varphi_{x_0}(\sigma(t)), \quad \sigma(t) \in T_{x_0}\mathcal{M}, \quad t \in [0, t_f],$$

for small enough  $t_f$ , see [63] for a precise definition.

In most applications of interest, see for example [5, 31],  $\mathcal{M}$  is a matrix manifold endowed with a Lie group action and there is a natural way to define a metric and a retraction. In fact, let  $\mathcal{M}$  be a manifold acted upon by a Lie group  $G$ , with a locally transitive group action  $\Lambda(g, x) = \Lambda_x(g)$ . Let us also consider a coordinate map  $\psi$ ,

$$\psi: \mathfrak{g} \rightarrow G, \text{ and } \rho_x := (\Lambda_x \circ \psi)'|_0. \quad (32)$$

One can prove that if there exists a linear map  $a_x: T_x\mathcal{M} \rightarrow \mathfrak{g}$  such that  $\rho_x \circ a_x = \text{Id}_{T_x\mathcal{M}}$ , then  $\varphi_x$ , given by

$$\varphi_x(v) := (\Lambda_x \circ \psi \circ a_x)(v), \quad (33)$$

is a retraction, see [16]. The existence of  $a_x$  is guaranteed, at least locally, by the transitivity of the action and the fact that  $\psi$  is a local diffeomorphism. The approach is analogous to the one described for differential equations in section 2.1. Therefore, we can construct retractions using any coordinate map from the Lie algebra,  $\mathfrak{g}$ , to the group.

Any metric on  $\mathfrak{g}$ ,  $\langle \cdot, \cdot \rangle_{\mathfrak{g}}$  induces a metric on  $\mathcal{M}$  by

$$\langle v_x, w_x \rangle_{\mathcal{M}} = \langle a_x(v_x), a_x(w_x) \rangle_{\mathfrak{g}}.$$

Also, we may define the image of the tangent space under the map  $a_x$ :

$$\mathfrak{m}_x := a_x(T_x\mathcal{M}) \subset \mathfrak{g}. \quad (34)$$

The set  $\mathfrak{m}_x$  is a linear subspace of the Lie algebra  $\mathfrak{g}$ , often of lower dimension. Parametrizations of the solution of (31) involving the whole Lie algebra are in general more computationally intensive than those restricted to  $\mathfrak{m}_x$ , but, if the isotropy is chosen suitably, they might lead to methods which converge faster to the optimum.

For the sake of illustration, we consider the minimization on a two-dimensional torus  $T^2 = S^1 \times S^1$ . Here we denote by  $S^1$  the circle, i.e.

$$S^1 = \{g \cdot \mathbf{e}_1 \mid \mathbf{R}^2, g(\alpha) \in SO(2)\},$$

$$g(\alpha) = \exp(\alpha E), \quad E = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad 0 \leq \alpha < 2\pi,$$

where  $\mathbf{e}_1$  is the first canonical vector and  $SO(2)$  is the commutative Lie group of planar rotations. Any element in  $T^2$  is of the form

$$x_0 \in T^2, \quad x_0 = (g(\theta)\mathbf{e}_1, g(\varphi)\mathbf{e}_1), \quad g(\theta), g(\varphi) \in SO(2).$$

The Lie group acting on  $T^2$  is  $SO(2) \times SO(2)$ , its corresponding Lie algebra is  $\mathfrak{so}(2) \times \mathfrak{so}(2)$ , which has dimension  $d = 2$  and basis  $\{(E, O), (O, E)\}$ , where  $O$  is the zero element in  $\mathfrak{so}(2)$ .

The Lie group action is

$$\Lambda_{x_0}(h_1, h_2) = (h_1 \cdot g(\theta)\mathbf{e}_1, h_1 \cdot g(\varphi)\mathbf{e}_1), \quad (h_1, h_2) \in SO(2) \times SO(2),$$

and  $\psi = \exp$ . Any  $v_{x_0} \in T_{x_0}T^2$  can be written as

$$v_{x_0} = (\alpha E\mathbf{e}_1, \beta E\mathbf{e}_1)$$

for some  $\alpha, \beta \in \mathbf{R}$ , so

$$a_{x_0}(v_{x_0}) = (\alpha E, \beta E).$$

Assume the cost function we want to minimize is simply the distance from a fixed plane in  $\mathbf{R}^3$ , say the plane with equation  $y = 8$ . This gives

$$\phi(g(\theta)\mathbf{e}_1, g(\varphi)\mathbf{e}_1) = |(1 + \cos(\theta)) \sin(\varphi) - 8|, \quad (35)$$

and the minimum is attained in  $\theta = 0$  and  $\varphi = \pi/2^3$ . In Figure 7 we plot  $-\text{grad } \phi$ , the negative gradient vector field for the given cost function. The Riemannian metric we used is,

$$\langle (\alpha_1 E\mathbf{e}_1, \beta_1 E\mathbf{e}_1), (\alpha_2 E\mathbf{e}_1, \beta_2 E\mathbf{e}_1) \rangle_{T^2} = \alpha_1 \alpha_2 + \beta_1 \beta_2,$$

and  $(\alpha_1 E\mathbf{e}_1, \beta_1 E\mathbf{e}_1) \in T_{(\mathbf{e}_1, \mathbf{e}_1)}T^2$ . This metric can be easily interpreted as a metric on the Lie algebra  $\mathfrak{g} = \mathfrak{so}(2) \times \mathfrak{so}(2)$ :

$$\langle (\alpha_1 E, \beta_1 E), (\alpha_2 E, \beta_2 E) \rangle_{\mathfrak{g}} = \alpha_1 \alpha_2 + \beta_1 \beta_2.$$

At the point  $p_0 = (g(\theta_0)\mathbf{e}_1, g(\varphi_0)\mathbf{e}_1) \in T^2 = S^1 \times S^1$  the gradient vector field can be represented by

$$(\gamma \cdot E \cdot g(\theta_0)\mathbf{e}_1, \delta \cdot E \cdot g(\varphi_0)\mathbf{e}_1),$$

where  $\gamma$  and  $\delta$  are real values given by

$$\gamma = -C \cdot \sin(\theta_0) \sin(\varphi_0), \quad \delta = C \cdot (1 + \cos(\theta_0)) \cos(\varphi_0),$$

and

$$C = 2((1 + \cos(\theta_0)) \sin(\varphi_0) - 8).$$

Gradient flows are not the only type of differential equations which can be used to solve optimization problems on manifolds. Alternative equations have been proposed in the context of neural networks [11, 12]. Often they arise naturally as the Euler–Lagrange equations of a variational problem.

---

<sup>3</sup>We have used a parameterization of  $T^2$  in  $\mathbf{R}^3$  in angular coordinates, obtained applying the following

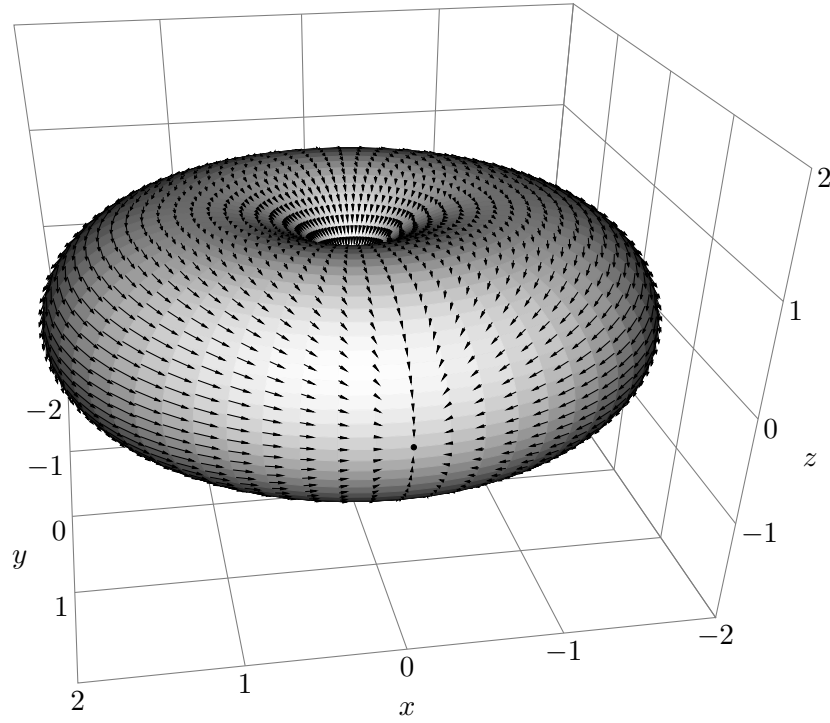


Figure 7: The gradient vector field of the cost function  $\phi((g(\theta)\mathbf{e}_1, g(\varphi)\mathbf{e}_1) = (1 + \mathbf{e}_1^T g(\theta)\mathbf{e}_1) \cdot \mathbf{e}_2^T g(\varphi)\mathbf{e}_1 - 8)^2$  on the torus. The vector field points towards the two minima, the global minimum is marked with a black spot in the middle of the picture.

## 6.2 Principal component analysis

Data reduction techniques are statistical signal processing methods that aim at providing efficient representations of data. A well-known data compression technique consists of mapping a high-dimensional data space into a lower dimensional representation space by means of a linear transformation. It requires the computation of the data covariance matrix and then the application of a numerical procedure to extract its eigenvalues and the corresponding eigenvectors. Compression is then obtained by representing the signal in a basis consisting only of those eigenvectors associated with the most significant eigenvalues.

In particular, Principal Component Analysis (PCA) is a second-order adaptive statistical data processing technique that helps removing the second-order correlation among given random signals. Let us consider a stationary multivariate random process  $x(t) \in \mathbf{R}^n$  and suppose its covariance matrix  $A = E[(x - E[x])(x - E[x])^T]$  exists and is bounded. Here the symbol  $E[\cdot]$  denotes statistical expectation. If  $A \in \mathbf{R}^{n \times n}$  is not diagonal, then the components of  $x(t)$  are statistically correlated. One can remove this redundancy by partially diagonalizing  $A$ , i.e. computing the operator  $F$  formed by the eigenvectors of the matrix  $A$  corresponding to its largest eigenvalues. This is possible since the covariance matrix  $A$  is symmetric (semi) positive-definite, and  $F \in \text{St}(n, p)$ . As a consequence the new random signal defined by  $y(t) := F^T(x(t) - E[x(t)]) \in \mathbf{R}^p$  has uncorrelated components, with  $p \leq n$  properly selected. The component signals of  $y(t)$  are the so called *principal components of the signal*  $x(t)$ , and their relevance is proportional to the corresponding eigenvalues  $\sigma_i^2 = E[y_i^2]$  which here are arranged in descending order ( $\sigma_i^2 \geq \sigma_{i+1}^2$ ).

Thus, the data stream  $y(t)$  is a compressed version of the data stream  $x(t)$ . After the reduced-size data has been processed (i.e. stored, transmitted), it needs to be recovered, that is, it needs to be brought back to the original structure. However, the principal-component-based data reduction technique is not lossless, thus only an approximation  $\hat{x}(t) \in \mathbf{R}^n$  of the original data stream may be recovered. As  $F$  is a tall-skinny orthogonal operator, an approximation of  $x(t)$  is given by  $\hat{x}(t) = Fy(t) + E[x]$ : Such approximate data stream minimizes the reconstruction error  $E[\|x - \hat{x}\|_2^2]$ , which equals  $\sum_{i=n+1}^p \sigma_i^2$ .

For a scalar or a vector-valued random variable  $x \in \mathbf{R}^n$  endowed with a probability density function  $p_x: x \in \mathbf{R}^n \rightarrow p_x(x) \in \mathbf{R}$ , the expectation of a function  $\beta: \mathbf{R}^n \rightarrow \mathbf{R}$  is defined as

$$E[\beta] := \int_{\mathbf{R}^n} \beta(x) p_x(x) d^n x. \quad (36)$$

Under the hypothesis that the signals whose expectation is to be computed are ergodic, the

---

mapping

$$(g(\theta)\mathbf{e}_1, g(\varphi)\mathbf{e}_1) \rightarrow \begin{cases} x &= (1 + \mathbf{e}_1^T g(\theta)\mathbf{e}_1) \cdot \mathbf{e}_1^T g(\varphi)\mathbf{e}_1 &= (1 + \cos(\theta)) \cos(\varphi), \\ y &= (1 + \mathbf{e}_1^T g(\theta)\mathbf{e}_1) \cdot \mathbf{e}_2^T g(\varphi)\mathbf{e}_1 &= (1 + \cos(\theta)) \sin(\varphi), \\ z &= \mathbf{e}_2^T g(\theta)\mathbf{e}_1 &= \sin(\theta), \end{cases}$$

with  $0 \leq \theta, \varphi < 2\pi$ . This is equivalent to the composition of two planar rotations and one translation in  $\mathbf{R}^3$ .

actual expectation (ensemble average) may be replaced by temporal-average on the basis of the available signal samples, namely

$$E[\beta] \approx \frac{1}{T} \sum_{t=1}^T \beta(x(t)). \quad (37)$$

### 6.3 Independent Component Analysis

An interesting example of a problem that can be tackled via statistical signal processing is the *cocktail-party problem*. Let us suppose  $n$  signals  $x_1(t), \dots, x_n(t)$  were recorded from  $n$  different positions in a room where there are  $p$  sources or speakers. Each recorded signal is a linear mixture of the voices of the sources  $s_1(t), \dots, s_p(t)$ , namely

$$x_1(t) = a_{1,1}s_1(t) + \dots + a_{1,p}s_p(t),$$

$$\vdots$$

$$x_n(t) = a_{n,1}s_1(t) + \dots + a_{n,p}s_p(t),$$

where the  $n \cdot p$  coefficients  $a_{i,j} \in \mathbf{R}$  denote the mixing proportions. The mixing matrix  $A = (a_{i,j})$  is unknown. The cocktail party problem consists in estimating signals  $s_1(t), \dots, s_p(t)$  from the only knowledge of their mixtures  $x_1(t), \dots, x_n(t)$ . The main assumption on the source signals is that  $s_1(t), \dots, s_p(t)$  are statistically independent. This problem can be solved using Independent Component Analysis.

Typically, one has  $n > p$ , namely, the number of observations exceeds the number of actual sources. Also, a typical assumption is that the source signals are spatially white, which means  $E[ss^T] = I_p$ . The aim of independent component analysis is to find estimates  $y(t)$  of signals in  $s(t)$  by constructing a de-mixing matrix  $W \in \mathbf{R}^{n \times p}$  and by computing  $y(t) := W^T x(t)$ . Using statistical signal processing methods, the problem is reformulated into an optimization problem on homogeneous manifolds for finding the de-mixing matrix  $W$ .

The geometrical structure of the parameter space in ICA comes from a signal pre-processing step named *signal whitening*, which is operated on the observable signal  $x(t) \rightarrow \tilde{x}(t) \in \mathbf{R}^p$  in such a way that the components of the signal  $\tilde{x}(t)$  are uncorrelated and have variances equal to 1, namely  $E[\tilde{x}\tilde{x}^T] = I_p$ . This also means that redundant observations are eliminated and the ICA problem is brought back to the smallest dimension  $p$ . This can be done by computing  $E[xx^T] = VDV^T$ , with  $V \in \text{St}(n, p)$  and  $D \in \mathbf{R}^{p \times p}$  diagonal invertible. Then pre-whitening is obtained as

$$\tilde{x}(t) := D^{-\frac{1}{2}} V^T x(t), \quad (38)$$

and  $E[\tilde{x}\tilde{x}^T] = \tilde{A}E[ss^T]\tilde{A}^T = \tilde{A}\tilde{A}^T = I_p$ .

After observable signal pre-whitening, the de-mixing matrix may be searched for such that it solves the optimization problem

$$\max_{W \in O(p)} \phi(W). \quad (39)$$

As explained, after pre-whitening, the number of projected observations in the signal  $\tilde{x}(t)$  equals the number of sources. However, in some applications it is known that not all the source signals are useful, so it is sensible to analyze only a few of them. In these cases, if we denote by  $\bar{p} \ll p$  the actual number of independent components that are sought after, the appropriate way to cast the optimization problem for ICA is

$$\max_{W \in \text{St}(n, \bar{p})} \phi(W), \text{ with } \bar{p} \ll p. \quad (40)$$

As a possible principle for reconstruction, the maximization or minimization of non-Gaussianity is viable. It is based on the notion that the sum of independent random variables has distribution closer to Gaussian than the distributions of the original random variables. A measure of non-Gaussianity is the kurtosis, defined for a scalar signal  $z \in \mathbf{R}$  as

$$\text{kurt}(z) := E[z^4] - 3E^2[z^2]. \quad (41)$$

If the random signal  $z$  has unitary variance, then the kurtosis computes as  $\text{kurt}(z) = E[z^4] - 3$ . Maximizing or minimizing kurtosis is thus a possible way of estimating independent components from their linear mixtures, see [12] and references therein for more details.

## 6.4 Computation of Lyapunov exponents

The Lyapunov exponents of a continuous dynamical system  $x' = F(x)$ ,  $(x(t) \in \mathbb{R}^n)$  provide a qualitative measure of its complexity. They are numbers related to the linearization  $A(t)$  of  $x' = F(x)$  along a trajectory  $x(t)$ . Consider the solution  $U$  of the matrix problem

$$\dot{U} = A(t)U, \quad U(0) = U_0, \quad n \times n.$$

The logarithms of the eigenvalues of the matrix

$$\Lambda = \lim_{t \rightarrow \infty} (U(t)^T U(t))^{\frac{1}{2t}},$$

are the Lyapunov exponents for the given dynamical system. In [22] a procedure for computing just  $k$  of the  $n$  Lyapunov exponents of a dynamical system is presented. The exponents are computed by solving an initial value problem on  $\text{St}(n, k)$  and computing a quadrature of the diagonal entries of a  $k \times k$  matrix-valued function. The initial value problem is defined as follows:

$$\dot{Q} = (A - QQ^T A + QSQ^T)Q,$$



with random initial value in  $\text{St}(n, k)$  and

$$S_{k,j} = \begin{cases} (Q^T A Q)_{k,j}, & k > j, \\ 0, & k = j, \\ -(Q^T A Q)_{j,k}, & k < j, \end{cases} \quad k, j = 1, \dots, p.$$

It can be shown that the  $i$ -th Lyapunov exponent  $\lambda_i$  can be obtained as

$$\lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t B_{i,i}(s) ds, \quad i = 1, \dots, k, \quad (42)$$

and

$$B = Q^T A Q - S.$$

One could use for example the trapezoidal rule to approximate the integral (42) and compute  $\lambda_i$  ( $i = 1, \dots, k$ ). We refer to [22] for further details on the method, and to [14] for the use of Lie group integrators on this problem. Lie group methods for ODEs on Stiefel manifolds have also been considered in [16, 35, 13].

We have here presented a selection of applications of Lie group integrators with particular emphasis to problems of evolution in classical mechanics and problems of signal processing. This is by no means an exhaustive survey; other interesting areas of application are for example problems in vision and medical imaging.

## Acknowledgment

This research was supported by a Marie Curie International Research Staff Exchange Scheme Fellowship within the 7th European Community Framework Programme. The authors would like to acknowledge the support from the GeNuIn Applications and SpadeAce projects funded by the Research Council of Norway, and most of the work was carried out while the authors were visiting Massey University, Palmerston North, New Zealand and La Trobe University, Melbourne, Australia.

## References

- [1] K. Baarman. *Direct minimization and acceleration of electronic structure calculations*. PhD thesis, Aalto University, Helsinki, Finland, 2012.
- [2] S. Blanes, F. Casas, J. A. Oteo J., and Ros. The Magnus expansion and some of its applications. *Phys. Rep.*, 470(5-6):151–238, 2009.

- [3] N. Bou-Rabee and J. E. Marsden. Hamilton-Pontryagin integrators on Lie groups. I. Introduction and structure-preserving properties. *Found. Comput. Math.*, 9(2):197–219, 2009.
- [4] T. J. Bridges and S. Reich. Multi-symplectic integrators: numerical schemes for Hamiltonian PDEs that conserve symplecticity. *Phys. Lett. A*, 284:184–193, 2001.
- [5] R. Brockett. Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems. *Lin. Alg. Applic.*, 146:79–91, 1991.
- [6] M. P. Calvo, A. Iserles, and A. Zanna. Runge-Kutta methods for orthogonal and isospectral flows. *Appl. Numer. Math.*, 22:153–163, 1996.
- [7] M. P. Calvo, A. Iserles, and A. Zanna. Numerical solution of isospectral flows. *Math. Comp.*, 66(220):1461–1486, 1997.
- [8] F. Casas and B. Owren. Cost Efficient Lie Group Integrators in the RKMK Class. *BIT Numerical Mathematics*, 43(4):723–742, 2003.
- [9] E. Celledoni. A note on the numerical integration of the KdV equation via isospectral deformations. *J. Phys. A: Math. Gen.*, 34(11):2205–2214, 2001.
- [10] E. Celledoni, F. Fassò, N. Säfström, and A. Zanna. The exact computation of the free rigid body motion and its use in splitting methods. *SIAM J. Sci. Comp.*, 30:2084–2112, 2008.
- [11] E. Celledoni and S. Fiori. Neural learning by geometric integration of reduced ‘rigid-body’ equations. *J. Comput. Appl. Math.*, 172:247–269, 2004.
- [12] E. Celledoni and S. Fiori. Descent methods for optimization on homogeneous manifolds. *Math. Comput. Simulation*, 79(4):1298–1323, 2008.
- [13] E. Celledoni, A. Marthinsen, and B. Owren. Commutator-free Lie group methods. *Future Generation Computer Systems*, 19:341–352, 2003.
- [14] E. Celledoni and B. Owren. A class of low complexity intrinsic schemes for orthogonal integration. *SIAM J. Numer. Anal.*, 40(6):2069–2084 (2003), 2002.
- [15] E. Celledoni and B. Owren. Lie group methods for rigid body dynamics and time integration on manifolds. *Comput. Methods Appl. Mech. Engrg.*, 192:421–438, 2003.
- [16] E. Celledoni and B. Owren. On the implementation of Lie group methods on the Stiefel manifold. *Numer. Algorithms*, 32(2-4):163–183, 2003.
- [17] E. Celledoni and N. Säfström. Efficient time-symmetric simulation of torqued rigid bodies using Jacobi elliptic functions. *Journal of Physics A*, 39:5463–5478, 2006.

- [18] E. Celledoni and N. Säfström. A Hamiltonian and multi-Hamiltonian formulation of a rod model using quaternions. *Comput. Methods Appl. Mech. Engrg.*, 199:2813–2819, 2010.
- [19] J. Cortès and S. Martínez. Non-holonomic integrators. *Nonlinearity*, 14:1365–1392, 2001.
- [20] P. E. Crouch and R. Grossman. Numerical integration of ordinary differential equations on manifolds. *J. Nonlinear Sci.*, 3:1–33, 1993.
- [21] I. Degani, A. Zanna, L. Sælen, and R. Nepstad. Quantum control with piecewise constant control functions. *SIAM J. Sci. Comput.*, 31(5):3566–3594, 2009.
- [22] L. Dieci and E. S. Van Vleck. Computation of a few Lyapunov exponents for continuous and discrete dynamical systems. *Applied Numerical Mathematics*, 17:275–291, 1995.
- [23] F. Diele, L. Lopez, and R. Peluso. The Cayley transform in the numerical solution of unitary differential systems. *Adv. Comput. Math.*, 8(4):317–334, 1998.
- [24] K. Engø. Partitioned Runge-Kutta methods in Lie-group setting. *BIT*, 43(1):21–39, 2003.
- [25] K. Engø and S. Faltinsen. Numerical integration of Lie–Poisson systems while preserving coadjoint orbits and energy. *SIAM J. Numer. Anal.*, 39(1):128–145, 2001.
- [26] J. Frank. Geometric space-time integration of ferromagnetic materials. *Appl. Numer. Math.*, 48(3-4):307–322, 2004. Workshop on Innovative Time Integrators for PDEs.
- [27] O. Gonzalez. Time integration and discrete Hamiltonian systems. *J. Nonlinear Sci.*, 6:449–467, 1996.
- [28] E. Hairer, Ch. Lubich, and G. Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010. Structure-preserving algorithms for ordinary differential equations, Reprint of the second (2006) edition.
- [29] E. Hairer and G. Vilmart. Preprocessed discrete Moser-Veselov algorithm for the full dynamics of a rigid body. *J. Phys. A*, 39:13225–13235, 2006.
- [30] F. Hausdorff. Die symbolische Exponential Formel in der Gruppen theorie. *Leipziger Ber.*, 58:19–48, 1906.
- [31] U. Helmke and J. B. Moore. *Optimization and dynamical systems.*, volume 14 of *Communications and Control Engineering Series*. Springer-Verlag London, Ltd., London, first edition edition, 1994.

- [32] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numer.*, 19:209–286, 2010.
- [33] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna. Lie-group methods. *Acta Numerica*, 9:215–365, 2000.
- [34] G.A. Jensen, N. Säfström, T.D. Nguyen, and T.I. Fossen. A nonlinear PDE formulation for offshore vessel pipeline installation. *Ocean. Engrg.*, 37(4):365–377, 2010.
- [35] S. Krogstad. A low complexity Lie group method on the Stiefel manifold. *BIT*, 43(1):107–122, 2003.
- [36] W. Kutta. Beitrag zur näherungsweisen Integration totaler Differentialgleichungen. *Zeitschr. für Math. u. Phys.*, 46:435–453, 1901.
- [37] J. Laskar, P. Robutel, F. Joutel, Gastineau, A.C.M. Correia, and B. Levrard. A long-term numerical solution for the insolation quantities of the earth. *Astron. Astrophys.*, 428:261–285, 2004.
- [38] T. Lee, M. Leok, and N. H. McClamroch. Lie group variational integrators for the full body problem. *Comput. Methods Appl. Mech. Engrg.*, 196(29-30):2907–2924, 2007.
- [39] B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*, volume 14 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, first edition edition, 2004.
- [40] M. Leok, T. Lee, and N.H. McClamroch. A Lie group variational integrator for the attitude dynamics of a rigid body with applications to the 3D pendulum. In *Proc. IEEE Conf. on Control Applications*, pages 962–967., 2005.
- [41] D. Lewis and P. J. Olver. Geometric integration algorithms on homogeneous manifolds. *Found. Comput. Math.*, 2(4):363–392, 2002.
- [42] D. Lewis and J. C. Simo. Conserving algorithms for the dynamics of Hamiltonian systems of Lie groups. *J. Nonlinear Sci.*, 4:253–299, 1994.
- [43] S. Leyendecker, P. Betsch, and P. Steinmann. Objective energy-momentum conserving integration for the constrained dynamics of geometrically exact beams. *Comput. Methods Appl. Mech. Engrg.*, 195:2313–2333, 2006.
- [44] A. Lundervold and H. Munthe-Kaas. Hopf algebras of formal diffeomorphisms and numerical integration on manifolds. In *Combinatorics and physics*, volume 539 of *Contemp. Math.*, pages 295–324. Amer. Math. Soc., Providence, RI, 2011.
- [45] J. E. Marsden and M. West. Discrete mechanics and variational integrators. *Acta Numerica*, 10:357–514, 2001.

- [46] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux. Geometric integration using discrete gradients. *Phil. Trans. Royal Soc. A*, 357:1021–1046, 1999.
- [47] R. I. McLachlan and C. Scovel. A survey of open problems in symplectic integration. *Integration algorithms and classical mechanics (Toronto, ON, 1993)*, 10:151–180, 1993.
- [48] R. I. McLachlan and A. Zanna. The discrete Moser–Veselov algorithm for the free rigid body, revisited. *Found. of Comp. Math.*, 5(1):87–123, 2005.
- [49] R.I. McLachlan and M.Perlmutter. Integrators for nonholonomic mechanical systems. *J. Nonlinear Sci.*, 16:283–328, 2006.
- [50] H. Munthe-Kaas. Lie–Butcher theory for Runge–Kutta methods. *BIT*, 35(4):572–587, 1995.
- [51] H. Munthe-Kaas. Runge–Kutta methods on Lie groups. *BIT*, 38(1):92–111, 1998.
- [52] H. Munthe-Kaas. High order Runge–Kutta methods on manifolds. *Appl. Numer. Math.*, 29:115–127, 1999.
- [53] H. Munthe-Kaas and B. Owren. Computations in a free Lie algebra. *Phil. Trans. Royal Soc. A*, 357:957–981, 1999.
- [54] H. Munthe-Kaas, G. R. W. Quispel, and A. Zanna. Application of symmetric spaces and Lie triple systems in numerical analysis. <http://www.ii.uib.no/~anto/PAPERS/muntheKaas01aos.pdf>, 2001.
- [55] H. Munthe-Kaas and A. Zanna. Numerical integration of differential equations on homogeneous manifolds. In F. Cucker and M. Shub, editors, *Foundations of Computational Mathematics*, pages 305–315. Springer Verlag, 1997.
- [56] R.M. Murray and S. S. Sastry. Nonholonomic motion planning: steering using sinusoids. *IEEE Trans. Automat. Control*, 38:700–716, 1993.
- [57] B. Owren and A. Marthinsen. Runge–Kutta methods adapted to manifolds and based on rigid frames. *BIT*, 39(1):116–142, 1999.
- [58] B. Owren and A. Marthinsen. Integration methods based on canonical coordinates of the second kind. *Numer. Math.*, 87:763–790, 2001.
- [59] G. R. W. Quispel and D. I. McLaren. A new class of energy-preserving numerical integration methods. *J. Phys. A*, 41(4):045206, 7, 2008.
- [60] I. Romero and F. Armero. An objective finite element approximation of the kinematics of geometrically exact rods and its use in the formulation of an energy-momentum conserving scheme in dynamics. *Internat. J. Numer. Methods Engrg.*, 54(12):1683–1716, 2002.

- [61] A. Saccon. Midpoint rule for variational integrators on Lie groups. *Internat. J. Numer. Methods Engrg.*, 78(11):1345–1364, 2009.
- [62] N. P. Säfström. *Modeling and simulation of rigid body and rod dynamics via geometric methods*. PhD thesis, Norwegian University of Science and Technology, Trondheim, Norway, 2009.
- [63] M. Shub. Some remarks on dynamical systems and numerical analysis. In L. Lara-Carrero and J. Lewowicz, editors, *Dynamical systems and partial differential equations*, volume VII ELAM, pages 62–92, 1986.
- [64] J. C. Simo. A finite strain beam formulation. the three-dimensional dynamic problem. part I. *Comput. Methods Appl. Mech. Engrg.*, 49:55–70, 1985.
- [65] J. C. Simo, N. Tarnow, and M. Doblaré. Non-linear dynamics of threedimensional rods: exact energy and momentum conserving algorithms. *Internat. J. Numer. Methods Engrg.*, 38:1431–1473, 1995.
- [66] J. C. Simo and L. Vu-Quoc. On the dynamics in space of rods undergoing large motions—a geometrically exact approach. *Comput. Methods Appl. Mech. Engrg.*, 66:125–161, 1988.